

ASSISTANT COMMISSIONER FOR PATENTS

WASHINGTON, D.C. 20231

ATTENTION: BOX PATENT APPLICATION

Sir:

Transmitted herewith for filing is the patent application of

Inventor(s): **Judith W. Zyskind**

For: **CHITOBIASE AS A REPORTER ENZYME**

Enclosed are:

- (X) 5 sheet(s) of drawing.
- (X) Return prepaid postcard.
- (X) Sequence Listing in 18 pages
- (X) Sequence Submission Form
- (X) Sequence Listing in Computer Readable Form

CLAIMS AS FILED

FOR	NUMBER FILED	NUMBER EXTRA	RATE	FEE
Basic Fee			\$690	\$690
Total Claims	50 - 20 =	30 ×	\$18	\$540
Independent Claims	12 - 3 =	9 ×	\$78	\$702

**TOTAL FILING FEE** \$1,932

- (X) A check in the amount of \$1,932 to cover the filing fee is enclosed.
- (X) The Commissioner is hereby authorized to charge any additional fees which may be required, now or in the future, or credit any overpayment to Account No. 11-1410. A duplicate copy of this sheet is enclosed.
- (X) Please use Customer No. 20,995 for the correspondence address.

*Daniel Hart*

Daniel Hart

Registration No. 40,637

Attorney of Record

S:\DOCS\DOH\DOH-4664.DOC  
080100



08/02/00

LOUIS J. KNOBBE  
DON W. MARTENS\*  
GORDON H. OLSON\*  
JAMES B. BEAR  
DARRELL L. OLSON\*  
WILLIAM B. BUNKER  
WILLIAM H. NIEMAN  
ARTHUR S. ROSE  
JAMES F. LESNIAK  
NED A. ISRAELSEN  
DREW S. HAMILTON  
JERRY T. SEWELL  
JOHN B. SGANGA, JR.  
EDWARD A. SCHLATTER  
GERARD VON HOFFMANN  
JOSEPH R. RE  
CATHERINE J. HOLLAND  
JOHN M. CARSON  
KAREN VOGEL WEIL  
ANDREW H. SIMPSON  
JEFFREY L. VAN HOOSEAR  
DANIEL E. ALTMAN  
MARGUERITE L. GUNN  
STEPHEN C. JENSEN  
VITO A. CANUSO III  
WILLIAM H. SHREVE  
LYNDA J. ZADRA-SYMES\*  
STEVEN J. NATAUPSKY  
PAUL A. STEWART  
JOSEPH F. JENNINGS  
CRAIG S. SUMMERS  
ANNEMARIE KAISER  
BRENTON R. BABCOCK

THOMAS F. SMEGAL, JR.  
MICHAEL H. TRENHOLM  
DIANE M. REED  
JONATHAN A. BARNEY  
RONALD J. SCHOENBAUM  
JOHN R. KING  
FREDERICK S. BERRETTA  
NANCY WAYS VENSKO  
JOHN P. GIEZENTANNER  
ADEEL S. AKHTAR  
GINGER R. DREGER  
THOMAS R. ARNO  
DAVID N. WEISS  
DANIEL HART, PH.D.  
DOUGLAS G. MUEHLHAUSER  
LORI LEE YAMATO  
MICHAEL K. FRIEDLAND  
STEPHEN M. LOBBIN  
STACEY R. HALPERN  
DALE C. HUNT, PH.D.  
LEE W. HENDERSON, PH.D.  
DEBORAH S. SHEPHERD  
RICHARD E. CAMPBELL  
MARK M. ABUMERI  
JON W. GURKA  
ERIC M. NELSON  
MARK R. BENEDICT, PH.D.  
PAUL N. CONOVER  
ROBERT J. ROBY  
SABING H. LEE  
KAROLINE A. DELANEY  
JOHN W. HOLCOMB  
JAMES J. MULLEN, III, PH.D.

# INTELLECTUAL PROPERTY LAW

## KNOBBE, MARTENS, OLSON & BEAR

A LIMITED LIABILITY PARTNERSHIP INCLUDING  
PROFESSIONAL CORPORATIONS

### PATENT, TRADEMARK AND COPYRIGHT CAUSES

501 WEST BROADWAY

SUITE 1400

SAN DIEGO, CALIFORNIA 92101-3505

(619) 235-8550

FAX (619) 235-0176

INTERNET: WWW.KNOB.COM

JOSEPH S. CIANFRANI  
JOSEPH M. REISMAN, PH.D.  
WILLIAM R. ZIMMERMAN  
GLEN L. NUTTALL  
ERIC S. FURMAN, PH.D.  
TIRZAH ABE LOWE  
GEOFFREY Y. IIDA  
ALEXANDER S. FRANCO  
SANJIVPAL S. GILL  
SUSAN M. MOSS  
JAMES W. HILL, M.D.  
ROSE M. THIESSEN, PH.D.  
MICHAEL L. FULLER  
MICHAEL A. GUILIANA  
MARK J. KERTZ  
RABINDER N. NARULA  
BRUCE S. ITCHKAWITZ, PH.D.  
PETER M. MIDGLEY  
THOMAS S. MCCLENAHAN  
MICHAEL S. OKAMOTO  
JOHN M. GROVER  
MALLARY K. DE MERLIER  
IRFAN A. LATEEF  
AMY C. CHRISTENSEN  
SHARON S. NG  
MARK J. GALLAGHER, PH.D.  
DAVID G. JANKOWSKI, PH.D.  
BRIAN C. HORNE  
PAYSON J. LEMEILLEUR  
WILLIAM G. BERRY  
DIANA W. PRINCE

#### OF COUNSEL

JERRY R. SEILER  
PAUL C. STEINHARDT

JAPANESE PATENT ATTY  
KATSUHIRO ARAI\*\*

EUROPEAN PATENT ATTY  
MARTIN HELLEBRANDT

KOREAN PATENT ATTY  
MINCHEOL KIM

SCIENTISTS & ENGINEERS  
(NON-LAWYERS)

RAIMOND J. SALENIKIS\*\*  
NEIL S. BARTFELD, PH.D.\*\*  
DANIEL E. JOHNSON, PH.D.\*\*  
JEFFERY KOEPKE, PH.D.\*\*  
KHURRAM RAHMAN, PH.D.  
JENNIFER A. HAYNES, PH.D.  
BRENDAN P. O'NEILL, PH.D.  
THOMAS Y. NAGATA  
LINDA H. LIU  
YASHWANT VAISHNAV, PH.D.  
MEGUMI TANAKA  
CHE S. CHERESKIN, PH.D.\*\*  
ERIK W. ARCHBOLD  
PHILIP C. HARTSTEIN  
JULIE A. HOPPER  
CHRIS S. CASTLE  
JAMES W. AUSLEY  
R. P. CARON, PH.D.  
JENNIFER HAYES  
KIRK E. PASTORIAN, PH.D.  
CHARLES T. RIDGELY  
KEITH R. MCCOLLUM  
LANG J. MCHARDY

\* A PROFESSIONAL CORPORATION  
\* ALSO BARRISTER AT LAW (U.K.)  
\*\* U.S. PATENT AGENT

Assistant Commissioner for Patents  
Washington, D.C. 20231

## CERTIFICATE OF MAILING BY "EXPRESS MAIL"

Attorney Docket No. : ELITRA.006A  
Applicant(s) : Judith W. Zyskind  
For : CHITOBIASE AS A REPORTER ENZYME  
Attorney : Daniel Hart  
"Express Mail"  
Mailing Label No. : EL 702 680 132 US  
Date of Deposit : August 2, 2000

I hereby certify that the accompanying

Transmittal in Duplicate; Specification in 31 pages; 5 sheets of drawings; 18 pages of sequence listing; sequence submission statement; sequence listing in computer readable form; Check for Filing Fee; Return Prepaid Postcard

are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and are addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

John Riedel

S:\DOCS\DOH\DOH-4662.DOC  
080100

201 CALIFORNIA STREET  
SUITE 1150  
SAN FRANCISCO, CALIFORNIA 94111  
(415) 954-4114  
FAX (415) 954-4111

620 NEWPORT CENTER DRIVE  
SIXTEENTH FLOOR  
NEWPORT BEACH, CALIFORNIA 92660  
(949) 760-0404  
FAX (949) 760-9502

3801 UNIVERSITY AVENUE  
SUITE 710  
RIVERSIDE, CALIFORNIA 92501  
(909) 781-9231  
FAX (909) 781-4507

1900 AVENUE OF THE STARS  
SUITE 1425  
LOS ANGELES, CALIFORNIA 90067  
(310) 551-3450  
FAX (310) 551-3459

**CHITOBIASE AS A REPORTER ENZYME**  
**GOVERNMENT INTEREST IN THE INVENTION**

Work performed in developing the present invention was supported, in part, by  
5 National Science Foundation grant MCB-9507209, NIH MBRS grant GM45765, and  
NIH NIGMS MARC F31 GM14967-0451. Accordingly, the United States Government  
may have certain rights in the invention.

Related Applications

This application claims priority to U.S. Provisional Patent Application Serial  
10 Number 60/159,221, filed October 13, 1999, the disclosure of which is incorporated  
herein by reference in its entirety.

Background of the Invention

Field of the Invention

15 Reporter enzymes are enzymes whose activities are easily assayed when present  
inside cells. In order to study the regulation of a gene whose expression is regulated by  
various environmental and/or cellular factors or influences, a gene encoding a reporter  
enzyme may be fused to the coding region or to the regulatory region of the regulated  
gene.

20 One of the most popular cytoplasmic reporter enzymes for use in bacteria is  $\beta$ -  
galactosidase. It is widely used in the art; however, because bacteria such as  
*Escherichia coli* contain an endogenous  $\beta$ -galactosidase encoding gene and  $\beta$ -  
galactosidase may be present in the cytoplasm of such bacteria, deletions of the LacZ  
gene, the source of the enzyme, must be introduced into the host cell line prior to its  
25 use. One goal of the present invention is to provide an alternative intracellular enzyme  
for use as a reporter.

This invention relates to genetic constructs and methods of using a cytoplasmic  
form of the chitobiase enzyme as a reporter. The invention also comprises expression  
vectors which express the cytoplasmic form of chitobiase. As used herein, all instances  
30 of the terminology "chitobiase" refer to a form of chitobiase which is present in the  
cytoplasm of the cell. Cytoplasmic forms of chitobiase may be generated via genetic  
engineering or microbial selection techniques.

### Summary of the Invention

One embodiment of the present invention is a method for characterizing a promoter comprising providing a construct comprising the promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase, introducing the construct into host cells, and identifying sequences in the promoter which regulate transcription levels. In one aspect of this embodiment, the cytoplasmic form of chitobiase lacks a signal sequence. In another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein which comprises a cytoplasmic form of chitobiase fused to a heterologous polypeptide. In still another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of *Alteromonas sp. 0-7*, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas sp. S9*, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*. In a further aspect of this embodiment, the method of identifying sequences which are involved in directing transcription comprises mutagenizing the promoter. In another aspect of this embodiment, the method of identifying sequences which are involved in transcription comprises constructing deletions in the promoter.

Another embodiment of the present invention is a method for identifying a regulatory element capable of directing or regulating transcription within a test nucleic acid sequence comprising providing a construct comprising the test nucleic acid sequence operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase, introducing the construct into host cells, and determining the level of chitobiase activity. In one aspect of this embodiment, the cytoplasmic form of chitobiase lacks a signal sequence. In another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, the fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide. In still another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form encodes a

cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of *Alteromonas* sp. 0-7, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*,  
5 *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas* sp. S9, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*. In one aspect of this embodiment, the reporter gene construct is  
10 introduced transiently. In another aspect of this embodiment, the reporter gene construct is introduced stably. The host cells may be selected from the group consisting of prokaryotic cells and eukaryotic cells. In another aspect of this embodiment, the method further comprises permeabilizing or lysing the host cells. The permeabilizing or lysing step may comprise treating the host cells with toluene. The step of  
15 determining the level of chitobiase activity may be selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate. In another embodiment, the step of  
20 determining the level of chitobiase activity may comprise determining the level of *p*-nitrophenol released from a substrate. In another aspect of this embodiment, the test nucleic acid sequence comprises a portion of genomic DNA. In a further aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after exposing the host cells to a desired set  
25 of environmental conditions. In still another aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after contacting the host cells with a compound to be tested for its influence on the level of transcription from said regulatory element.

Another embodiment of the present invention is a method of detecting  
30 successful transformation, comprising the steps of introducing a nucleic acid encoding a cytoplasmic form of chitobiase into host cells, and detecting chitobiase expression in the host cells.

Another embodiment of the present invention is a fusion protein-reporter gene construct comprising a promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitinase fused in frame with a nucleic acid encoding a heterologous polypeptide, wherein the heterologous polypeptide is not  $\beta$ -galactosidase or a portion thereof, and wherein the heterologous polypeptide does not contain a signal peptide. In one aspect of this embodiment, the nucleic acid encodes a cytoplasmic form of chitinase obtained from an organism selected from the group consisting of *Alteromonas* sp. 0-7, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas* sp. S9, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*. In another aspect of this embodiment, the nucleic acid further comprises a  $\lambda$  site-specific recombination sequence.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pJMF3.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pJMF4.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pDYK9.

Another embodiment of the present invention is a reporter gene construct comprising plasmid pDYK11.

Another embodiment of the present invention is a host cell comprising a fusion protein-reporter gene construct comprising a promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitinase fused in frame with a nucleic acid encoding a heterologous polypeptide, wherein the heterologous polypeptide is not  $\beta$ -galactosidase or a portion thereof, and wherein the heterologous polypeptide does not contain a signal peptide. In one aspect of this embodiment, the nucleic acid is integrated into a chromosome of the cell. In another aspect of this embodiment, the nucleic acid is transiently expressed in the host cell.

Another embodiment of the present invention is a nucleic acid encoding a cytoplasmic form of chitobiase in which the signal sequence of native chitobiase has been inactivated or deleted. In one aspect of this embodiment, the signal sequence has been mutated to inactivate it.

5           An isolated or purified polypeptide comprising a cytoplasmic form of chitobiase fused in frame with a heterologous polypeptide, wherein the heterologous polypeptide is not  $\beta$ -galactosidase or a portion thereof and wherein the heterologous polypeptide does not contain a signal peptide.

10           Another embodiment of the present invention is an isolated or purified polypeptide comprising a cytoplasmic form of chitobiase in which the signal peptide of native chitobiase has been inactivated or deleted. In one aspect of this embodiment, the signal sequence has been mutated to inactivate it.

15           Another embodiment of the present invention is a method for monitoring the activity of a promoter comprising providing a construct comprising the promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase, introducing the construct into host cells, and determining the level of chitobiase activity. In one aspect of this embodiment, the cytoplasmic form of chitobiase lacks a signal sequence. In another aspect of this embodiment, the nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, the fusion protein comprising  
20           a cytoplasmic form of chitobiase fused to a heterologous polypeptide. In one aspect of this embodiment, the nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of *Alteromonas sp. 0-7*, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba*  
25           *histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas sp. S9*, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*. In another  
30           aspect of this embodiment, the reporter gene construct is introduced transiently. In still another aspect of this embodiment, the reporter gene construct is introduced stably. In a further aspect of this embodiment, the host cells are selected from the group consisting

of prokaryotic cells and eukaryotic cells. In yet another aspect of this embodiment, the method further comprises permeabilizing or lysing the host cells. For example, the permeabilizing or lysing step may comprise treating the host cells with toluene. In yet another aspect of this embodiment, the step of determining the level of chitobiase activity may be selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate. In still another aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of *p*-nitrophenol released from a substrate. In still another aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after exposing the host cells to a desired set of environmental conditions. In a further aspect of this embodiment, the step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after contacting the host cells with a compound to be tested for its influence on the level of transcription from said regulatory element. For example, the compound may comprise a compound to be tested for activity as a drug.

#### Brief Description of the Drawings

Figure 1 illustrates plasmids pJMF3 and pJMF4 containing *attP* in 2 different orientations and the *lac* promoter with the first 21 amino acids of *lacZα* (from pUC19) fused in-frame to the *chb* gene. The sequence of the fusion region is shown in Figure 2. Restriction sites shown are found once in the plasmid sequences except for *NotI*, which has 2 sites flanking the P15A origin.

Figure 2 illustrates the sequence of the *lac* promoter and the chitobiase fusion found in pJMF3 and pJMF4 (SEQ ID NOS: 15 and 16). Fusion between *lacZα* (from pUC19) and *chb* [Soto-Gill, R.W. et al, J. Biol. Chem. 264:14778-14782 (1998)] is indicated by (/); start of transcription is indicated by (+1). Sequence and binding sites in the *lac* promoter regulatory region are found in Dickson, R.C. et al., Science 187:27-32 (1975), the disclosure of which is incorporated herein by reference in its entirety. Restriction enzyme sites shown in Figure 2 are found once in the plasmid sequences except for *SphI*, which has 2 sites; these different restriction sites can be used to replace



the *lac* promoter with another promoter together with part of a coding region to produce an in-frame fusion with *chb*.

Figure 3 depicts integration of *chb* fusions into the chromosome by site-specific recombination between *attB* and *attP*. The steps involved are described below and in Diederich, L.L, et al., Plasmid 28:14-24 (1992), the disclosure of which is incorporated herein by reference in its entirety.

Figure 4 illustrates plasmids pDYK9 and pDYK11 which contain *dnaA-chb* fusions. pDYK9 is deleted for the *rpmH* regulatory region. The orientation of *attP* in pDYK9 and pDYK11 is the same as that of pJMF3. After integration at *attB* of the larger *NotI* fragment, the *dnaA* promoters are oriented to transcribe in the same direction as replication fork movement.

Figure 5 illustrates the *rmpH-dnaA* regulatory region and *dnaA*-chitobiase fusion (SEQ ID NO: 17). The *dnaA* box and promoters are shaded, and the coding region of the *dnaA* gene is striped. The fusion contains two amino acids (between the backslashes) from pUC19. The region cloned into pDYK9 is between *SphI* primer II and *KpnI* primer, and the region cloned into pDYK11 is between *SphI* primer I and *KpnI* primer. The numbers above the primers refer to the nucleotide in the sequence amplified. For sequence numbering and the locations of promoters and protein binding sites, see Froelich, J.M. et al., J. Bacteriol. 178:6006-6012 (1996); Hansen, F.G. et al., EMBO J. 1:1043-1048 (1982) and Messer, W. and C.W. Weigel, Initiation of Chromosome Replication. p. 1579-1601 (1996) In F.C. Neidhart, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter and H.E. Umbarger (Eds.), Escherichia coli and Salmonella Cellular and Molecular Biology., ASM Press, Washington, D.C., the disclosures of which are incorporated herein by reference in their entirety.

#### Detailed Description of the Preferred Embodiment

Reporter genes and reporter gene constructs play a number of important roles in a variety of molecular biology techniques. For example, reporter genes may be used to determine whether a sequence contains a promoter or other cis-acting element which directs transcription, such as an enhancer. In addition, reporter genes may be used to identify regulatory sites in promoters or other cis-acting elements and to determine the

effects of mutating these regulatory sites on the level of gene expression directed by the promoters or other cis-acting elements. Reporter genes may also be used to detect successful transformation. In addition, reporter genes may be used to monitor gene expression under various conditions and to identify drugs.

5           The structure of a reporter gene construct containing the cytoplasmic form of chitobiase will vary according to its purposes. When the reporter gene construct is a vector, one must decide between a vector that incorporates the reporter gene into the host's genome or one that replicates extrachromosomally, such as a plasmid. When the reporter construct is not designed to integrate into the host genome, the vector can  
10       contain an origin of replication with activity in the host cell of interest. This feature provides the reporter gene vector the ability to replicate within the host cell in which it has been introduced. In addition to the origin of replication, reporter constructs often contain a promoter, a multiple cloning site, a selectable marker, and of course a reporter gene. Reporter constructs for use in eukaryotic cells may also contain a polyA site  
15       adjacent to the reporter gene.

          Given the utility of reporter gene constructs, it is not surprising that a number of cytoplasmic reporter gene constructs and different reporter genes are available for use by those of skill in the art. For example, the cytoplasmic reporter enzymes  
20       chloramphenicol acetyltransferase (CAT), firefly luciferase,  $\beta$ -glucuronidase (GUS), green fluorescent protein (GFP), and  $\beta$ -galactosidase have been used extensively. However, such reporters all have individual shortcomings that may limit or preclude their usage under some conditions. For example, high levels of GFP are toxic to the cell. In addition, reporter enzymes are not expressed equally in all cell types nor are they equally stable when expressed in all cell types. Furthermore, there is a recognized  
25       need for multiple reporter enzymes that can be assayed independently of one another in order to simultaneously study the regulation of multiple genes within a single cell type. Therefore, there exists a continuing need to identify reporter enzymes with useful properties.

          The cytoplasmic enzyme  $\beta$ -galactosidase is widely used as a reporter gene in  
30       various microbiological and molecular biological studies. This enzyme is used in both *in vitro* and *in vivo* assays. The wide acceptance of this reporter system results, in part, because it is non-isotopic and extremely flexible. It is used in a number of assay

formats and has an extremely broad linear range. Nevertheless, because  $\beta$ -galactosidase is present in the cytoplasm of various host cells such as *Escherichia coli*, deletion of the *lacZ* gene, the source of the enzyme, is often required prior to its use in a host cell system. One goal of the present invention was to provide an alternative intracellular enzyme for use as a reporter.

An extensive discussion of various molecular biology techniques is available in Ausubel, et al., (eds) "Short Protocols in Molecular Biology," Wiley and Sons, Inc., New York (1997), the disclosures of which are incorporated herein by reference in their entirety. Examples of such techniques include isolating and preparing DNA for manipulation, gel electrophoresis, polymerase chain reaction (PCR), determining nucleic acid sequences, screening nucleic acid libraries, mutagenesis of DNA, and introducing DNA into host cells.

The structure of a reporter gene construct will vary according to its purposes. The reporter gene constructs are constructed according to standard techniques of molecular biology well known in the art.

When the reporter gene construct is a vector, one must decide between a vector that incorporates the reporter gene into the host's genome or one that replicates extrachromosomally, such as a plasmid. When integration of the reporter gene is a desired result, a reporter gene construct will contain sequences that will facilitate incorporation.

One example of integration sequences that can be included in a reporter gene construct is the  $\lambda$  *attP* site. This site permits a single copy of the reporter gene to be incorporated into a host bacterial genome. Integration-promoting sequences with utility in mammalian cells include the long terminal repeats found in retroviral genomes. These sequences promote viral genome integration in a host genome and have been used extensively by those of skill in the art to promote the integration of exogenous sequences in mammalian host cells.

When the reporter construct is not designed to integrate into the host genome, it is common that the vector contain an origin of replication with activity in the host cell of interest. This feature provides the reporter gene vector the ability to replicate within the host cell in which it has been introduced.

In addition to the origin of replication, reporter constructs will often contain additional features that promote the expression of the nucleic acid sequence or sequences contained in the construct. These additional sequences can include a polyA site, a multiple cloning site, a drug resistance marker, and of course a reporter gene.

5           The present invention relates to the use of chitobiase as a reporter gene. The chitobiase may be used as a reporter in bacteria, plants, mammalian cells and other host cell lines. One possible alternative to using  $\beta$ -galactosidase as a reporter gene was to develop cytoplasmic N,N'-diacetylchitobiase (*N*-acetyl- $\beta$ -D-glucosaminidase, EC 3.2.1.30) for use as a reporter enzyme. One advantage of the enzyme *N*-acetyl- $\beta$ -D-  
10   glucosaminidase or "chitobiase" over  $\beta$ -galactosidase is that genes encoding chitobiase are missing from many bacteria, including *E. coli*, some fungi, and some eukaryotic cells. Thus, it is not necessary to engineer many host cells to lack reporter activity as is the case with  $\beta$ -galactosidase. The present invention also relates to various protein expression vectors that can be used to express the reporter gene. In addition, the present  
15   invention may be used in conjunction with other reporter enzymes in systems in which the regulation or activities of multiple genes is to be studied simultaneously.

Chitobiase is one of two enzymes that hydrolyze chitin, an abundant insoluble polysaccharide, to its monomeric unit, *N*-acetylglucosamine (GlcNac). Chitobiase is known to be present in a number of organisms. For example, the chitobiase enzyme is  
20   known to be present in various genera including *Arabidopsis*, *Bacillus*, *Bombyx*, *Bos*, *Caenorhabditis*, *Candida*, *Dictyostelium*, *Entamoeba*, *Felis*, *Homo*, *Korat*, *Lactobacillus*, *Leishmania*, *Mus*, *Pisum*, *Porphyromonas*, *Pseudoalteromonas*, *Rattus*, *Serratia*, *Streptomyces*, *Sus*, *Trichoderma*, and *Vibrio*. Specific examples of organisms known to contain chitobiase include *Alteromonas* sp. 0-7, *Arabidopsis thaliana*,  
25   *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas* sp. S9, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*,  
30   *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*.

One source of the enzyme is the marine bacterium, *Vibrio harveyi*. *Escherichia coli* cells harboring a plasmid carrying the *chb* gene from *Vibrio harveyi* were reported to produce the enzyme, which was found to be associated with the outer membrane of the bacterial cells. These are described in Jannatipour, M. et al., "Translocation of *Vibrio harveyi* *N,N'*-diacetylchitobiase to the outer membrane of *Escherichia coli*," J. Bacteriol. 169:3785-3791 (1987) and Soto-Gil & Zyskind, *N,N'*-Diacetylchitobiase of *Vibrio harveyi* primary structure, processing, and evolutionary relationships. J. Biol Chem. 264:14778-14782 (1989), both of which are hereby incorporated by reference.

The present invention contemplates expressed cytoplasmic forms of chitobiase in various forms. In one embodiment, the signal sequence is deleted from the amino terminal portion of the protein. Presumably the removal of this sequence results in the expression of a cytoplasmic form of the enzyme that is not secreted from the host cell or incorporated into the membrane of the host cell producing the enzyme.

The present invention also contemplates the generation of fusion proteins comprising a fusion polypeptide joined in frame to chitobiase. Preferably, the fusion polypeptide comprises a polypeptide other than chitobiase, such as a heterologous protein. The heterologous polypeptide may comprise a polypeptide having a biological activity (such as an enzymatic or other activity besides activity as an immunogen) or the heterologous polypeptide may not have a biological activity. The heterologous polypeptide does not include a signal sequence which directs its secretion. Preferably the heterologous polypeptide is not  $\beta$ -galactosidase or a portion thereof. Thus, the fusion reporter gene construct contains a sequence encoding the fusion polypeptide genetically fused in frame with a sequence encoding chitobiase. In one embodiment, this fusion may remove the amino-terminal signal peptide sequence of chitobiase and replace it with a heterologous protein.

In another embodiment, the fusion protein construct comprises a chitobiase gene sequence that has been truncated to remove at least the signal peptide sequence of the gene. Alternatively, mutations may be introduced into the signal peptide sequence so that it is no longer functional. Such mutations may be introduced using a variety of techniques familiar to those skilled in the art, including site directed mutagenesis, cassette mutagenesis, and chemical mutagenesis.

Once the reporter gene construct is made it is introduced into a host cell line for testing. Host cells of prokaryotic and eukaryotic origin can be used with the reporter gene constructs of the present invention. A variety of methods are available to introduce reporter gene constructs into host prokaryotic cells. For example, bacteria can be transformed using calcium chloride, electroporation, or viral vectors such as the filamentous phages. These and other prokaryotic transformation protocols are well known in the art.

Alternatively, the sequence encoding chitobiase may be introduced in eukaryotic cells, including yeast, mammalian, plant, and insect cells. For example, the sequence encoding chitobiase may be inserted into a yeast artificial chromosome, a yeast plasmid, a bovine papilloma virus vector or other extrachromosomal vector, a retroviral vector, a Ti-plasmid, or a baculovirus vector. A variety of such vectors are known to those skilled in the art. The vectors may be introduced into any of the yeast, mammalian, plant, and insect cells familiar to those skilled in the art.

The introduction of the reporter gene construct into mammalian cells can likewise utilize a number of transfection protocols well known to those of skill in the art. As discussed above, transfections can be transient or stable. Examples of suitable transfer protocols include calcium phosphate transfection, DEAE-Dextran, electroporation, liposome-mediated transfection, and viral transfection. These and other eukaryotic transformation protocols are well known in the art.

Following introduction of the reporter gene construct into the host cell of interest, the enzymatic activity of the reporter gene is measured. Preferably, the chitobiase assays are performed after permeabilizing or lysing the host cells. There are a variety of cell permeabilization and cell lysis procedures available to those of ordinary skill in the art, including methods such as sonication or lysozyme treatment. One embodiment of the present invention uses toluene treatment to permeabilize cells. The details of this method are discussed in D.Y. Kalabat et al., *BioTechniques* 25:1030-1035 (1998) and Miller, J.H. *A Short Course in Bacterial Genetics*, CSH Laboratory Press, Cold Spring Harbor, NY 1992, the disclosures of which are incorporated herein by reference in their entirety.

Cellular chitobiase activity can be measured quantitatively by following the hydrolysis of chitobiase substrates. Examples of substrates with utility in chitobiase

activity assays include *N,N'*-diacetylchitobiose (chitobiose), *p*-nitrophenyl-*N*-acetyl- $\beta$ -D-glucosaminide (PNAG)(Sigma Chemical, St. Louis, MO), and 5-bromo-4-chloro-3-indolyl-*N*-acetyl- $\beta$ -D-glucosaminide (X-Gluc)(Sigma Chemical, St. Louis, MO). Other substrates are also contemplated for use in the assays of the present invention.

5           Products produced by the hydrolysis of the chitobiase substrates are monitored using various means familiar to those skilled in the art. For example, various optical means are known to those skilled in the art. One such optical means may comprise detection of chemiluminescent or fluorescent products released from a substrate. Alternatively, the level of chitobiase activity may be determined by measuring the  
10           amount of light absorbed by a product produced from a substrate or measuring a decrease in the amount of a detectable substrate. In one embodiment, *p*-nitrophenol is released from the substrate and measured at 400 nm. Other monitoring methods well known in the art can be used to quantitate signals produced in the chitobiase assay. These may include use of radioactive substrates or substrates having radiofrequency  
15           tags. In another embodiment, blue/white colony indicator plates are used to monitor enzyme activity.

          Another embodiment of the present invention is a kit. One aspect of this embodiment includes a reporter gene construct comprising a vector containing a  
20           chitobiase reporter gene. The reporter gene construct also contains a multiple cloning site containing a variety of restriction endonuclease cutting sites that facilitate the introduction of exogenous DNA into the construct.

          The kit embodiment of the present invention also includes those components necessary to assay for chitobiase activity produced by the reporter gene construct. For  
25           example, in one embodiment, the kit will include a supply of a suitable chitobiase substrate whose metabolism into product by the reporter enzyme can be assayed.

#### EXAMPLES

          The following Examples are disclosed to assist in the understanding of the present invention. The Examples below should not be construed to limit the scope of  
30           the invention and such variations of the invention now known or later developed, which would be within the purview of one of ordinary skill in the art and are considered to fall within the scope of the invention hereinafter claimed.

## Example 1

### Construction of Vectors

Vectors were constructed using polymerase chain reaction (PCR) products that were cloned first at the *Sma*I or *Eco*RV restriction sites of plasmid pBluescript II (pKSII+)(Stratagene; San Diego, CA). The nucleotide sequence of all PCR products was determined using standard techniques well known in the art. Unmethylated plasmid DNA was isolated from an *E. coli dam*<sup>-</sup> strain when cutting with the *Bcl*I enzyme was required.

The pDYK9 (SEQ ID NO: 11) plasmid was constructed by ligating a *Sph*I-*Kpn*I PCR product containing the *dnaA* promoter region into plasmid pRSG196 that contains the *V. harveyi* chitobiase gene (see Jannatipour, M., et al.). Briefly, the *V. harveyi* chitobiase gene was prepared by cloning a *Sau*3A partial digest of *V. harveyi* DNA into the single *Bam*HI site with the *tet* gene of pMK2004 [See Soto-Gill & Zyskind, "Cloning of *Vibrio harveyi* chitinase and chitobiase genes in *Escherichia coli*," pp. 209-223, In J.P. Zikakis (ed.), Chitin, chitosan, and related enzymes," Academic Press, Inc., New York (1984)]. The clones containing the gene of interest were detected by the presence of the yellow *p*-nitrophenylate product after individual colonies of the clone bank were sprayed with 10 mM PNAG dissolved in 100 mM sodium phosphate, pH 7.0. Restriction maps of the positive clones were made according to standard techniques well known in the art.

The *chb* gene was subcloned in a 3.5-kb *Eco*RI fragment from pRSG14 [See Jannatipour, M. et al., "Translocation of *Vibrio harveyi* *N,N'*-diacetylchitobiase to the outer membrane of *Escherichia coli*," J. Bacteriol. 169:3785-3791 (1987)] into the *Eco*RI site of pUC19 [Yanisch-Perron, c.J. et al., Gene 33: 103-119 (1985)]. Plasmid pRSG196 was constructed by deleting a 0.5-kb *Sst*I fragment from one of these clones [See Jannatipour, M. et al., "Translocation of *Vibrio harveyi* *N,N'*-diacetylchitobiase to the outer membrane of *Escherichia coli*," J. Bacteriol. 169:3785-3791 (1987)].

The pRSG196 plasmid was also cut with *Sph*I and *Kpn*I to accommodate the insertion of the *dnaA* fragment. Plasmid pAC17 [described in Chiaramello & Zyskind, "Coupling of DNA replication to growth rate in *Escherichia coli*: a possible role for guanosine tetraphosphate," J. Bacteriol. 172:2013-2019 (1992) hereby incorporated by reference] served as template with primers 5'-GCA CAT GCA TGC TGG TCA TTA



AAT TTT CC-3' (SEQ ID NO 1) and 5'-CGG GGT ACC AAC TCA TCC TGC AAT  
 CG-3' (SEQ ID NO 2) producing a PCR product 374 bp long that contains 353 bp from  
 the *dnaA* promoter region [bases 583 to 935, numbering according to Hansen, F. G., et  
 al., "The nucleotide sequence of the *dnaA* gene promoter and of the adjacent *rpmH*  
 5 gene, coding for the ribosomal protein L34, of *Escherichia coli*," EMBO J. 1:1043-1048  
 (1982)]. The forward primer (*SphI* primer II) contained an *SphI* site and the reverse  
 primer (*KpnI* primer) contained a *KpnI* site for cloning. This created an in-frame fusion  
 between the amino-terminal 17 amino acids of DnaA and the carboxy-terminal end of  
 chitobiase deleted for the amino-terminal 22 amino acids including the signal peptide.

10 The next step in constructing the pDYK9 vector was taking the 3270 bp *DraI*-  
*HindIII* (partial digest) fragment containing the *dnaA-chb* fusion and ligating it to the  
 pACYC184 2555 bp *HincII-HindIII* fragment carrying chloramphenicol-resistance  
 (Cm<sup>r</sup>) and the P15A origin. This fragment was described in Chang & Cohen,  
 "Construction and characterization of amplifiable multicopy DNA cloning vehicles  
 15 derived from the p15A cryptic miniplasmid," J. Bacteriol. 134:1141-1156 (1987).

A *NotI* site was introduced at the *AccI* site after digestion with *AccI*, treatment  
 with Mung bean nuclease, and ligation to phosphorylated *NotI* linkers (New England  
 Biolabs, Inc.; Beverly, MA). A *NotI* site was introduced at the *AseI* site after digestion  
 with *AseI*, treatment with Mung bean nuclease, and ligation to phosphorylated *NotI*  
 20 linkers (New England Biolabs, Inc.; Beverly, MA). An *XbaI-SphI* PCR product  
 containing the *rrnBt112* terminator was ligated into this plasmid cut with the same  
 enzymes creating pDYK7. *E. coli* chromosomal DNA served as template with primers  
 5'-CTA GTC TAG ATG CCG AAC TCA GAA GTG A-3' (SEQ ID NO 3) and 5'-GCA  
 CAT GCA TGC GGG GGA TGG CTT GTA GAT-3' (SEQ ID NO 4) to produce a PCR  
 25 product 357 bp long that contains bases 6534 to 6869 from the *rrnB* operon [numbering  
 according to Brosius, J., et al., "Gene organization and primary structure of a ribosomal  
 RNA operon from *Escherichia coli*." J. Mol. Biol. 148:107-127 (1981)] and includes  
 the complex transcription termination region of this operon which is described in Orosz,  
 A., et al., "Analysis of the complex transcription termination region of the *Escherichia*  
 30 *coli rrnB* gene," Eur. J. Biochem. 201:653-659 (1991). The forward primer contained  
 an *XbaI* site and the reverse primer contained an *SphI* site.

A *BclI*-*SmaI* PCR product containing the  $\lambda$  *attP* site was ligated into pDYK7 cut  
 with *Tth111I*, treated with Mung bean nuclease, then digested with *BclI*. Plasmid  
 pHN894 [described in Goodman, S. D., et al., "Deformation of DNA during site-  
 specific recombination of bacteriophage lambda: Replacement of IHF protein by HU  
 5 protein or sequence-directed bends," Proc. Natl. Acad. USA 89:11910-11914 (1992)]  
 served as template with the forward primer 5'-CAT GAT CAT GCG ACA GGT TTG  
 ATG A-3' (SEQ ID NO 5) and the reverse primer 5'-GGG GGC GCC TAC CTT TCA  
 CGA G-3' (SEQ ID NO 6) producing a PCR product 466 bp long that contains the  $\lambda$   
*attP* site. The 466 bp PCR product containing the  $\lambda$  *attP* site was first cloned into the  
 10 *SmaI* site of the pKSII+ plasmid to produce pDYK8. The forward primer contains a  
*BclI* site and the reverse primer contains G's at the 5' end in order to recreate a *SmaI*  
 site when cloning the PCR fragment into a *SmaI* site. This PCR product includes bases  
 -211 to +241 from the center of the *attP* core and the sequence required for optimum  $\lambda$   
*attP* site integration as described in Nagaraja & Weisberg, "Specificity determinants in  
 15 the attachment sites of bacteriophages HK022 and  $\lambda$ ," J. Bacteriol. 172:6540-6550  
 (1990). The orientation of *attP* is such that when the fusion is integrated at *attB*, the  
 transcription direction of *dnaAp1* and *p2* promoters is the same as replication fork  
 movement mimicking the orientation at the wild type *dnaA* promoters.

Vector pDYK11 (SEQ ID NO: 12) was constructed by ligating an *SphI*-*KpnI*  
 20 PCR product containing the *rpmH-dnaA* promoter region into pDYK9 also digested  
 with *SphI* and *KpnI*. Plasmid pAC17, described in Chiaramello & Zyskind, (1992),  
 served as template with the primers 5'-CAT GCA TGC ATG AAA CGA TGG ACA  
 CC-3' (SEQ ID NO 7) and 5'-CGG GGT ACC AAC TCA TCC TGC AAT CG-3' (SEQ  
 ID NO 8) to produce a PCR product 616 bp long that contains 598 bp from the *rpmH*-  
 25 *dnaA* regulatory region [bases 338 to 935, numbering according to Hansen, F. G., et al.,  
 (1982)]. The forward primer (*SphI* primer I) contained an *SphI* site and the reverse  
 primer (*KpnI* primer) contained a *KpnI* site for cloning. SEQ ID NO: 18 provides the  
 complete coding sequence of the DNA encoding the *dnaA*/chitobiase fusion protein.  
 The complete sequence of the *dnaA*/chitobiase fusion protein is provided in SEQ ID  
 30 NO: 19.

The vector pJMF3 (SEQ ID NO: 13) was constructed by first ligating an *AseI*  
 linker (5'-CATTAATGCATG-3' (SEQ ID NO 9) self-hybridized) into the *SphI* site of

pDYK11. The pUC19, [described in Yanisch-Perron C., et al., "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors," Gene 33:103-19 (1985)] *AseI-KpnI* fragment containing the *lacPO*-polylinker region was ligated to this plasmid after digestion with *AseI* and *KpnI*. The resulting in-frame fusion between the amino-terminal 21 amino acids of the pUC19 *lacZ*( $\alpha$ ) peptide [Yanisch-Perron C., et al., 1985] and the carboxy-terminal end of chitobiase deleted for the amino-terminal 22 amino acids is identical to the protein fusion in pRSG196 [Jannatipour, M., et al., (1995)].

Vector pJMF4 (SEQ ID NO: 14 ) was constructed by ligating the *Bam*HI-*Eco*RV fragment of pDYK8 containing the  $\lambda$  *attP* site into pDYK7 which had been digested with *Tth*111I, treated with Mung bean nuclease, then digested with *Bcl*I, to create pTKP9. This reversed the orientation of the *attP* site relative to the *attP* site in pDYK9. The pDYK11 *Bsp*MII-*KpnI* fragment containing the *rpmH-dnaA* promoter region was then ligated into pTKP9 digested with the same enzymes to create pTKP11. This reversed the orientation of the *attP* site relative to the *attP* site in pDYK11. An *AseI* linker (5'-CATTAATGCATG-3' (SEQ ID NO 10) self-hybridized) was ligated into the *Sph*I site of pTKP11 to create pJMF2. The pUC19 [as described in Yanisch-Perron C., et al., (1985)] *AseI-KpnI* fragment containing the *lacPO*-polylinker region was ligated to pJMF2 cut with *AseI* and *KpnI* to create pJMF4.

## Example 2

### Site-specific Recombination

To move the chitobiase fusions in pDYK9 and pDYK11 to the *attB* site in the chromosome, *Not*I fragments from these plasmids were self-ligated and transformed or electroporated into strain WM2269 (DH5 $\alpha$  containing pLDR8) [See Zyskind, J. W. and S. I. Bernstein (1992) "Recombinant DNA Laboratory Manual," Academic Press, San Diego, CA]. Plasmid pLDR8, described in Diederich, L. L. J., et al., "New cloning vectors for integration into the  $\lambda$  attachment site *attB* of *Escherichia coli* chromosome," Plasmid 28:14-24 (1992), expresses integrase from the  $\lambda$  *P<sub>R</sub>* promoter and contains the  $\lambda$  *cl<sub>857</sub>* repressor gene, a kanamycin resistance gene, and a temperature-sensitive origin of replication. The transformed or electroporated cells were incubated at 42°C with shaking for 30 min then moved to 37°C for 1 h followed by selection on Luria broth

agar plates containing chloramphenicol (25 µg/ml) at 42°C. Transformants were screened for loss of kanamycin resistance and, therefore, loss of pLDR8.

### Example 3

#### Bacteriophage P1 Transduction

5 Transduction with P1 bacteriophage by the method of Zyskind & Bernstein, Recombinant DNA Laboratory Manual. Academic Press, San Diego, CA (1992) was used to construct strains and to confirm the chromosomal location of the *dnaA-chb* fusions. Cotransduction of Cm<sup>r</sup> (carried by the fusion) and *galK* (linked to *attB*) indicated that Cm<sup>r</sup> and *galK* are linked on the chromosomes of strains DYK9W,  
10 DYK9F, DYK11W, and DYK11F.

### Example 4

#### p-Nitrophenol Chitinase Activity Assay

Chitinase activity is located in the cytoplasm when its signal peptide is replaced by fusion with another peptide. Accordingly, chitinase assays are performed  
15 on permeabilized or lysed cells. A variety of permeabilization or cell lysis protocols are available to liberate the enzyme from within a cell population being tested. One such protocol involves toluene-treated cells washed once with M9 salts according to the method of Miller, J.H. A Short Course in Bacterial Genetics, CSH Laboratory Press, Cold Spring Harbor, NY 1992, which is hereby incorporated by reference. The  
20 toluenized cells are placed in a chitinase buffer (10 mM Tris-HCl, pH 8.0, and 0.5 M NaCl) and 666 µM PNAG, the chitinase enzyme substrate. NaCl is included because chitinase has approximately 80% of full activity in the absence of salt, with maximal activity occurring between 0.25 and 0.6 M NaCl. Toluenized cells (0.772 ml) are preincubated at 28°C and the reaction started with the addition of 0.228 ml PNAG (1  
25 mg/ml). After incubation at 28°C, the reaction is stopped by the addition of 1 ml of 1 M Tris base. The release of *p*-nitrophenol is measured at 400 nm and turbidity at 550 nm. *p*-Nitrophenol release is measured immediately at 400 nm with a molar absorptivity of  $1.8 \times 10^3$  liters mol<sup>-1</sup> cm<sup>-1</sup>. Units are calculated after subtracting the light scattering factor ( $1.5 \times OD_{550}$ ) from  $OD_{400}$  of the sample. The normalizing factor of 1.5 was  
30 determined previously by measuring the light scattering ratio of bacteria at  $OD_{400}$  and  $OD_{550}$ . One unit of chitinase activity is the amount of enzyme that catalyzes the formation of 1 pmol of *p*-nitrophenol per min at 28°C. For comparison to Miller units

of  $\beta$ -galactosidase [described in Miller, J. H., A Short Course in Bacterial Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1992)], the units are normalized to 1 ml of culture at  $OD_{450}=1$ .

#### Example 5

##### 5 Determination of Chitinase Activity in Cells Containing Vectors pJM3 and pJM4

(Figure 1) contain the *lacPO* promoter with the first 21 amino acids of *lacZa* (from pUC19) fused in-frame to the *chb* gene. These plasmids also contain the  $\lambda$  phage *attP* recombination site in different orientations, the gene encoding chloramphenicol acetyltransferase (*cat*), and a ribosomal terminator, *rrnBt12*, inserted upstream of the *lac-chb* fusion to prevent read-through from other promoters. The chitinase activity associated with these plasmids (Table 1) is high in the absence of IPTG because of titration of *lac* repressor expressed from a single copy chromosomal gene. Induction by IPTG is approximately 10-fold (Table 1).

**Table 1. Chitinase Activity of *lacZ-chb* Fusion<sup>a</sup>**

Plasmid	Chitinase <sup>b</sup> (U <sup>c</sup> )	
	-IPTG	+IPTG (1 mM)
pJM3 in DH5 $\alpha$	668 $\pm$ 45	9320 $\pm$ 347
pJM4 in DH5 $\alpha$	788 $\pm$ 44	7188 $\pm$ 477

<sup>a</sup> Overnight cultures were diluted 1:1000 into 50 mL prewarmed LB and grown to  $OD_{450} = 0.1$ . 1 mM IPTG was added to half of the culture, and growth continued to  $OD_{450} = 0.3$ .

<sup>b</sup> Triplicate samples were assayed. Mean chitinase activities are given with standard deviations.

<sup>c</sup> One unit of chitinase activity is 1 pmol of *p*-nitrophenol/min at 28°C. Units given for 1 mL of culture at  $OD_{450} = 1$ .

The *lac* promoter can be replaced with another promoter and a fusion protein created with chitinase by cutting with *SphI* or *AseI* and either *SalI*, *KpnI*, or *SstI* (Figure 2). Fusions created with these vectors can be moved to the chromosome by site specific recombination at the  $\lambda$  *attB* site to permit single copy analysis of the activity of the promoter. The protocol, as described more fully in Diederich, et al., "New cloning vectors for integration into the  $\lambda$  attachment site *attB* of *Escherichia coli* chromosome,"

Plasmid 28:14-24 (1992), involves two components, (i) a circular DNA containing the  $\lambda$  attachment site, *attP*, the promoter-*chb* gene fusion, and the *cat* gene, and (ii) a helper plasmid, pLDR8, which contains the *int* gene under the control of the temperature-sensitive repressor, *cI857* and a temperature-sensitive origin of replication. The plasmid is digested with *NotI* to remove the P15A origin, and the fragment containing the chitobiase fusion is self-ligated prior to transformation into cells containing pLDR8. Integration occurs by site-specific recombination between *attP* and *attB* (17.4 min on the *E. coli* chromosome, Figure 3).

#### Example 6

##### Use of the Chitobiase Reporter Enzyme to Study *dnaA* Gene regulation

Two plasmids, pDYK9 and pDYK11, discussed in the Examples above, were constructed to assess the regulation of the *dnaA* gene using chitobiase as a reporter enzyme. These plasmids differ by the absence of the *rpmH* regulatory region in pDYK9 (Figure 4). These fusions were moved from the plasmid to the chromosomal *attB* site for single copy analysis as described above. After transformation of strain WM2269 with the ligated DNA, integration occurred by site specific recombination between the *attP* and the *attB* sites. The orientation of *attP* in pDYK9 and pDYK11 is such that when the fusion is integrated at *attB*, the transcription direction of *dnaAp1* and *dnaAp2* promoters is the same as movement of the replication fork. This orientation is the same as at the *dnaA* wild type location. The genetic location of the fusions was confirmed in the  $\text{Cm}^r$  transformants by demonstrating cotransduction of  $\text{Cm}^r$  and *galK*.

The fusions created in strain WM2269 were moved by P1 transduction to MG1655, creating strain DYK9W with pDYK9 and strain DYK11W with pDYK11. Deletion of the *rpmH* promoters had very little effect (1.4-fold) on chitobiase activity (Table 2, compare lines 1 and 3)

**Table 2. Chitobiase and  $\beta$ -Gal Activities of *dnaA-chb* and *dnaA-lacZ* Fusions in WT and *fis* Mutant Backgrounds**

Strains	$\beta$ -Gal <sup>c</sup>	
	Chitobiase <sup>a</sup> (U <sup>b</sup> )	(Miller U <sup>d</sup> )
DYK11W <i>fis</i> <sup>+</sup>	30.0 $\pm$ 2.1	
DYK11F <i>fis</i> ::985	80.0 $\pm$ 1.2	
DYK9W <i>fis</i> <sup>+</sup>	44.0 $\pm$ 2.1	
DYK9F <i>fis</i> ::985	96.0 $\pm$ 3.5	
RB220 <i>fis</i> <sup>+</sup>		59.8 $\pm$ 7.9
TP220 <i>fis</i> ::767		115.2 $\pm$ 7.2
<sup>a</sup> Triplicate samples were assayed during exponential growth. Mean chitobiase activities are given with standard deviation. <sup>b</sup> One unit of chitobiase activity is 1 pmol of <i>p</i> -nitrophenol per min at 28°C. Units given for 1 mL of culture at OD <sub>450</sub> = 1. <sup>c</sup> Data from Reference 8. <sup>d</sup> Unit defined in Reference 13.		

5           Fis protein binds to a site in the *dnaAp2* promoter that covers the -35 sequence (Figure 5), and appears to be a repressor of DnaA expression. A fusion protein with  $\beta$ -galactosidase activity that is expressed from the *rpmH-dnaA* regulatory region has increased  $\beta$ -galactosidase activity (1.9-fold) in a *fis*<sup>-</sup> mutant when compared to Fis wild type cells [as described in Froelich, J. M., et al., "Fis binding in the *dnaA* operon promoter region," J. Bacteriol. 178:6006-6012 (1996), data shown in Table 3].

10           Similarly, the absence of Fis leads to a greater than 2-fold increase in chitobiase activity of the DnaA-chitobiase fusion protein for the DYK9F and DYK11F strains, comparable in extent to that observed with the *dnaA-lacZ* fusion strain, TP220 (Table 3).

15           In the reporter gene constructs discussed in the Example, all upstream transcriptional activity was prevented from entering the *chb* reporter gene. The plasmid vectors, pJMF3 and pJMF4, described in the Examples above, contained the *rrnbt1t2* terminator upstream of the promoter fusion, which prevented readthrough from chromosomal promoters near the insertion site. Only chitobiase activity originating from the promoters of interest was expressed.

The *attP* site in these vectors allowed integration at the chromosomal *attB* in a specific orientation, depending on the vector used. With these vectors, any chitobiase fusion involving an essential gene can be moved to the chromosome, thus permitting single copy analysis with a chromosomal orientation similar to the wild-type gene.

#### Example 7

##### Identification of Promoters in Test Sequences

A nucleic acid prospectively containing a promoter is inserted upstream of a nucleic acid encoding a cytoplasmic form of chitobiase as described above. For example, the nucleic acid prospectively containing a promoter may be inserted into a restriction site in a sequence containing a plurality of restriction sites, such as a polylinker, which is located upstream of the nucleic acid encoding chitobiase. The test sequence may comprise any nucleic acid to be tested for promoter activity. In one embodiment, the test sequence may comprise a genomic DNA sequence. For example, the genomic DNA sequence may be a randomly generated DNA fragment, such as a fragment generated using shotgun cloning techniques, a restriction fragment, or any other sequence.

The vectors containing the test sequence upstream of the nucleic acid encoding chitobiase are introduced into an appropriate host cell. The level of chitobiase activity is assayed and compared to the level obtained from a control vector which lacks an insert in the cloning site. The presence of an elevated expression level in cell containing the vector containing the insert with respect to the level in cells containing the control vector without the insert indicates the presence of a promoter in the insert.

In some embodiments, the activity of the promoter in the test sequence may be assayed after exposure of the host cells to conditions which may influence the level of transcription from the promoter. For example, the environment of the host cells may be altered to determine whether the transcription level is influenced by environmental factors, including factors such as temperature, pH, nutrients, or availability of oxygen. In such analyses, chitobiase levels are assayed under a variety of environmental conditions to determine the effects of the environmental conditions on transcription levels from the promoter. In addition, the activity of the promoters may be examined in the presence or absence of compounds to be tested for regulatory activity. For example, the activity of the promoters may be tested by determining the levels of chitobiase produced in the presence or absence of compounds to be tested for activity as drugs.



Promoter sequences within the test sequences may be further defined by constructing nested deletions in the test sequences using conventional techniques such as Exonuclease III digestion. The resulting deletion fragments can be inserted into the promoter reporter vector to determine whether the deletion has reduced or obliterated promoter activity as determined by measuring chitobiase activity in cells containing the deletion vectors. In this way, the boundaries of the promoters may be defined. If desired, potential individual regulatory sites within the promoter may be identified using techniques such as site directed mutagenesis or linker scanning to obliterate potential transcription factor binding sites within the promoter individually or in combination. The effects of these mutations on transcription levels may be determined by inserting the mutations into the cloning sites in the promoter reporter vectors and measuring the levels of chitobiase produced from the mutated promoters.

The activity of known promoters may also be monitored by operably linking them to a nucleic acid encoding a cytoplasmic form of chitobiase. The activity of the promoters may be analyzed under various environmental conditions as described above. In addition, the activity of the promoters may be analyzed in the presence or absence of compounds to be tested for the ability to affect transcription from the promoters. For example, the compounds may be tested for activity as drugs.

In some embodiments, the chitobiase reporter constructs may be used in systems for identifying compounds that modulate cell surface protein-mediated activity or compounds which modulate the activities of intracellular signaling systems. Techniques for using reporter genes to identify compounds which modulate cell surface protein-mediated activity have been described in U.S. Patent Number 5,401,629 and U.S. Patent Number 5,436,128, the disclosures of which are incorporated herein by reference in their entireties. Briefly, in such methods, a construct comprising a promoter operably linked to a nucleic acid encoding a reporter enzyme is introduced into cells which express the cell surface protein and cells which do not express the cell surface protein. Each of the cells are contacted with test compounds and the effects of these compounds on transcription levels is measured by determining the level of activity of the reporter enzyme. The level of expression of the reporter gene in cells expressing the cell surface protein is compared to the level in cells which do not express the cell surface protein to identify compounds that modulate cell surface protein activity.

Similarly, the chitobiase reporter constructs may be used to identify compounds which influence the activity of intracellular signaling pathways, such as cAMP-based or phosphorylation-based pathways. In such methods, a promoter which is activated via such pathways is operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase.

5 The cells are contacted with test compounds. Those compounds which activate the pathway to which the promoter responds will produce an enhanced level of chitobiase activity in the cells as compared to the level of chitobiase activity in control cells which have not been contacted with the test compound.

#### Example 8

##### Detecting Successful Transformation or Transfection Using Chitobiase

10 A vector comprising a sequence encoding a cytoplasmic form of chitobiase operably linked to a sequence capable of directing transcription of the chitobiase gene is introduced into a host cell. The host cells are contacted with a chitobiase substrate and those host cells which contain chitobiase activity are identified as cells which were  
15 successfully transformed or transfected. In some embodiments, a portion or replica of a colony may be lysed or permeabilized prior and the lysate or permeabilized cells may be contacted with the chitobiase substrate.

#### CONCLUSION

20 New gene reporter systems that use chitobiase have been described. Chitobiase has advantages over other reporter gene systems in that chitobiase is not found in many cell lines traditionally used in reporter gene systems.

25 Finally, the forgoing examples are not intended to limit the scope of the present invention, which is set forth in the following claims. In particular, various equivalents and substitutions will be recognized by those of ordinary skill in the art in view of the foregoing disclosure, and these are contemplated to be within the scope of the present invention. All references cited herein are incorporated herein by reference in their entireties.

WHAT IS CLAIMED IS:

1. A method for characterizing a promoter comprising:
  - providing a construct comprising said promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase;
  - 5 introducing the construct into host cells; and
  - identifying sequences in said promoter which regulate transcription levels.
2. The method of Claim 1, wherein said cytoplasmic form of chitobiase lacks a signal sequence.
- 10 3. The method of Claim 2, wherein said nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, said fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide.
4. The method of Claim 1, wherein said nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of *Alteromonas* sp. 0-7, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas* sp. S9, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*.
- 15 5. The method of Claim 1, wherein said method of identifying sequences which are involved in directing transcription comprises mutagenizing said promoter.
- 25 6. The method of Claim 1, wherein said method of identifying sequences which are involved in transcription comprises constructing deletions in said promoter.
7. A method for identifying a regulatory element capable of directing or regulating transcription within a test nucleic acid sequence comprising:
  - providing a construct comprising said test nucleic acid sequence
  - 30 operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase;
  - introducing said construct into host cells; and
  - determining the level of chitobiase activity.

8. The method of Claim 7, wherein said cytoplasmic form of chitobiase lacks a signal sequence.

9. The method of Claim 8, wherein said nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, said fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide.

10. The method of Claim 7, wherein said nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of *Alteromonas* sp. 0-7, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas* sp. S9, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*.

11. The method of Claim 7, wherein said reporter gene construct is introduced transiently.

12. The method of Claim 7, wherein said reporter gene construct is introduced stably.

13. The method of Claim 7, wherein said host cells are selected from the group consisting of prokaryotic cells and eukaryotic cells.

14. The method of Claim 7, further comprising permeabilizing or lysing said host cells.

15. The method of Claim 14, wherein said permeabilizing or lysing step comprises treating said host cells with toluene.

16. The method of Claim 7, wherein said step of determining the level of chitobiase activity is selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate.

17. The method of Claim 7, wherein said step of determining the level of chitobiase activity comprises determining the level of *p*-nitrophenol released from a substrate.

18. The method of Claim 7, wherein said test nucleic acid sequence comprises a portion of genomic DNA.

19. The method of Claim 7, wherein said step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after exposing said host cells to a desired set of environmental conditions.

20. The method of Claim 7, wherein said step of determining the level of chitobiase activity comprises determining the level of chitobiase activity after contacting said host cells with a compound to be tested for its influence on the level of transcription from said regulatory element.

21. A method of detecting successful transformation, comprising the steps of:  
introducing a nucleic acid encoding a cytoplasmic form of chitobiase into host cells; and

detecting chitobiase expression in said host cells.

22. A fusion protein-reporter gene construct comprising a promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase fused in frame with a nucleic acid encoding a heterologous polypeptide, wherein said heterologous polypeptide is not  $\beta$ -galactosidase or a portion thereof and wherein said heterologous polypeptide does not contain a signal peptide.

23. The nucleic acid of Claim 22, wherein said nucleic acid encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of *Alteromonas* sp. 0-7, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*, *Porphyromonas gingivalis*, *Pseudoalteromonas* sp. S9, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*.

24. The nucleic acid of Claim 22, further comprising a  $\lambda$  site-specific recombination sequence.

25. A reporter gene construct comprising plasmid pJMF3.

26. A reporter gene construct comprising plasmid pJMF4.

5 27. A reporter gene construct comprising plasmid pDYK9.

28. A reporter gene construct comprising plasmid pDYK11.

29. A host cell comprising the construct of Claim 22.

30. The host cell of Claim 29 wherein said nucleic acid is integrated into a chromosome of said cell.

10 31. The host cell of Claim 29, wherein said nucleic acid is transiently expressed in said host cell.

32. A nucleic acid encoding a cytoplasmic form of chitobiase in which the signal sequence of native chitobiase has been inactivated or deleted.

15 33. The nucleic acid of Claim 32, wherein the signal sequence has been mutated to inactivate it.

34. An isolated or purified polypeptide comprising a cytoplasmic form of chitobiase fused in frame with a heterologous polypeptide, wherein said heterologous polypeptide is not  $\beta$ -galactosidase or a portion thereof and wherein said heterologous polypeptide does not contain a signal peptide.

20 35. An isolated or purified polypeptide comprising a cytoplasmic form of chitobiase in which the signal peptide of native chitobiase has been inactivated or deleted.

36. The polypeptide of Claim 35, wherein the signal sequence has been mutated to inactivate it.

25 37. A method for monitoring the activity of a promoter comprising:

providing a construct comprising said promoter operably linked to a nucleic acid encoding a cytoplasmic form of chitobiase;

introducing said construct into host cells; and

determining the level of chitobiase activity.

30 38. The method of Claim 37, wherein said cytoplasmic form of chitobiase lacks a signal sequence.

39. The method of Claim 38, wherein said nucleic acid encoding a cytoplasmic form of chitobiase encodes a fusion protein, said fusion protein comprising a cytoplasmic form of chitobiase fused to a heterologous polypeptide.

5 40. The method of Claim 37, wherein said nucleic acid encoding a cytoplasmic form encodes a cytoplasmic form of chitobiase obtained from an organism selected from the group consisting of *Alteromonas sp. 0-7*, *Arabidopsis thaliana*, *Bacillus subtilis*, *Bombyx mori*, *Bos taurus*, *Caenorhabditis elegans*, *Candida albicans*, *Dictyostelium discoideum*, *Entamoeba histolytica*, *Felis catus*, *Homo sapiens*, *Korat cats*, *Lactobacillus casei*, *Leishmania donovani*, *Mus musculus*, *Pisum sativum*,  
10 *Porphyromonas gingivalis*, *Pseudoalteromonas sp. S9*, *Rattus norvegicus*, *Serratia marcescens*, *Streptomyces plicatus*, *Streptomyces thermoviolaceus*, *Sus scrofa*, *Trichoderma harzianum*, *Vibrio furnissii*, *Vibrio harveyi*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*.

15 41. The method of Claim 37, wherein said reporter gene construct is introduced transiently.

42. The method of Claim 37, wherein said reporter gene construct is introduced stably.

43. The method of Claim 37, wherein said host cells are selected from the group consisting of prokaryotic cells and eukaryotic cells.

20 44. The method of Claim 37, further comprising permeabilizing or lysing said host cells.

45. The method of Claim 44, wherein said permeabilizing or lysing step comprises treating said host cells with toluene.

25 46. The method of Claim 37, wherein the step of determining the level of chitobiase activity is selected from the group consisting of measuring the amount of a chemiluminescent product produced from a substrate, measuring the amount of a fluorescent product produced from a substrate, measuring the amount of light absorbed by a product produced from a substrate and measuring a decrease in the amount of a detectable substrate.

30 47. The method of Claim 37, wherein said step of determining the level of chitobiase activity comprises determining the level of *p*-nitrophenol released from a substrate.





### Abstract of the Disclosure

The present invention relates to reporter gene constructs encoding a cytoplasmic form of chitinase (N,N'-diacetylchitinase) and methods of using these reporter gene constructs. The use of a cytoplasmic form of chitinase as a reporter enzyme is generally applicable to the study of gene expression in organisms which do not contain N-acetyl- $\beta$ -D-glucosaminidases.

10

S:\DOCS\DOH\DOH-4653.DOC  
072800

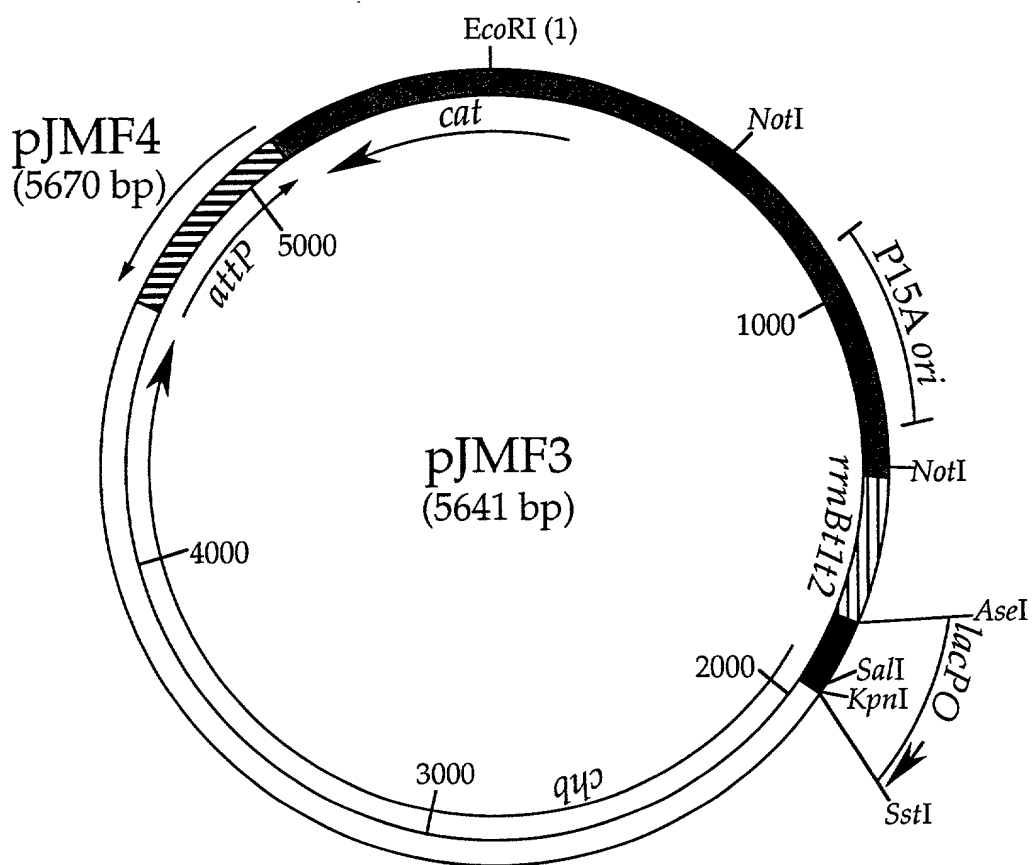


Figure 1

SEQ ID NO: 15:

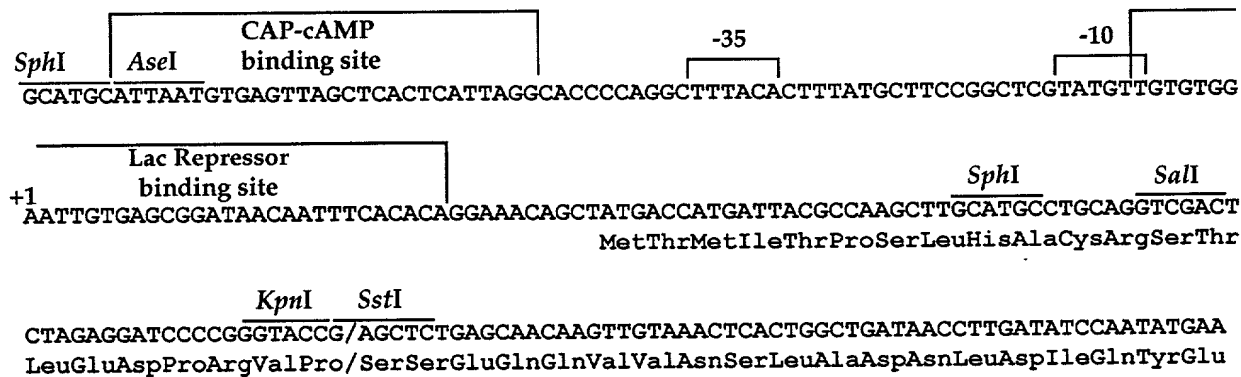


Figure 2

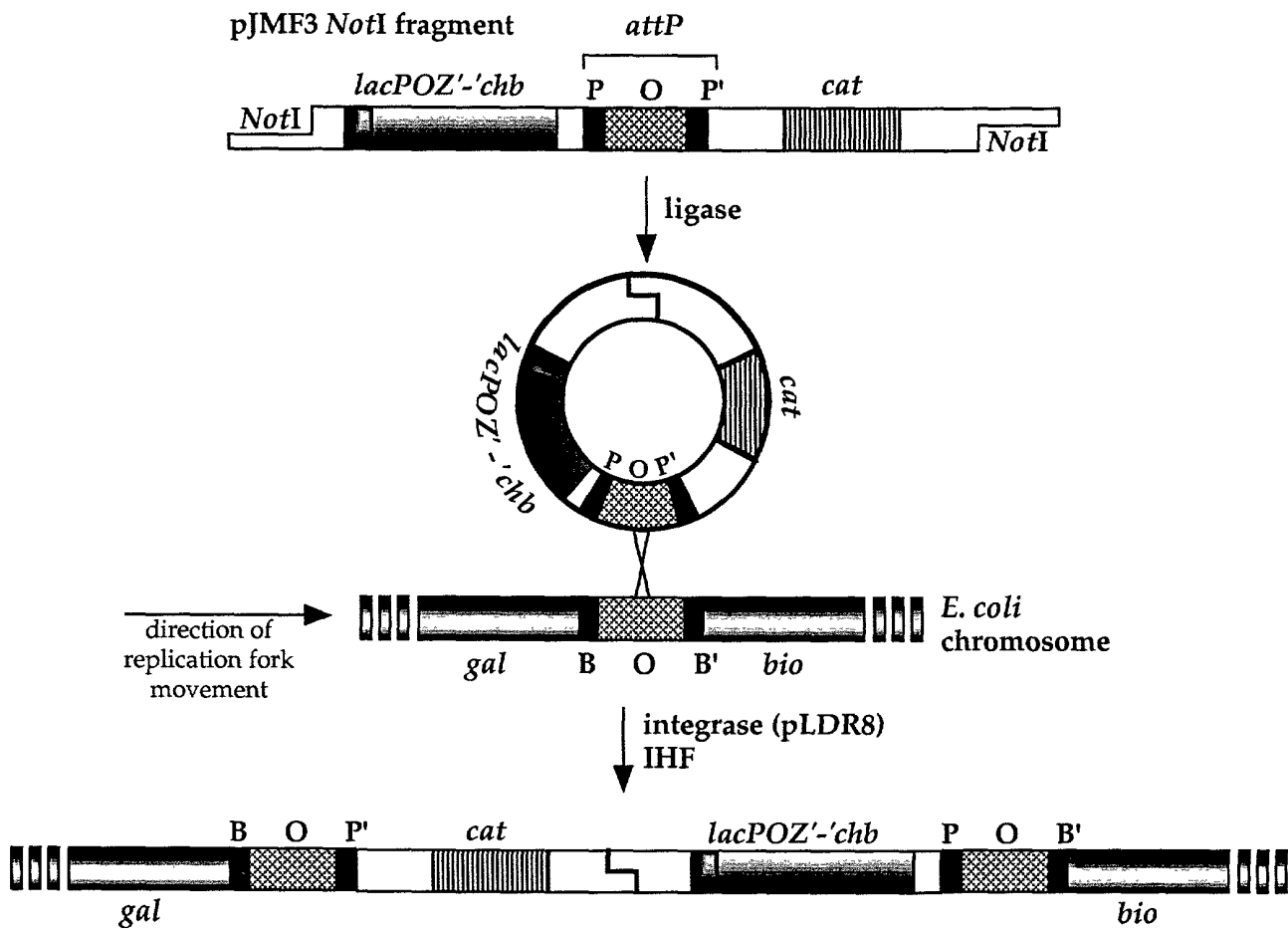
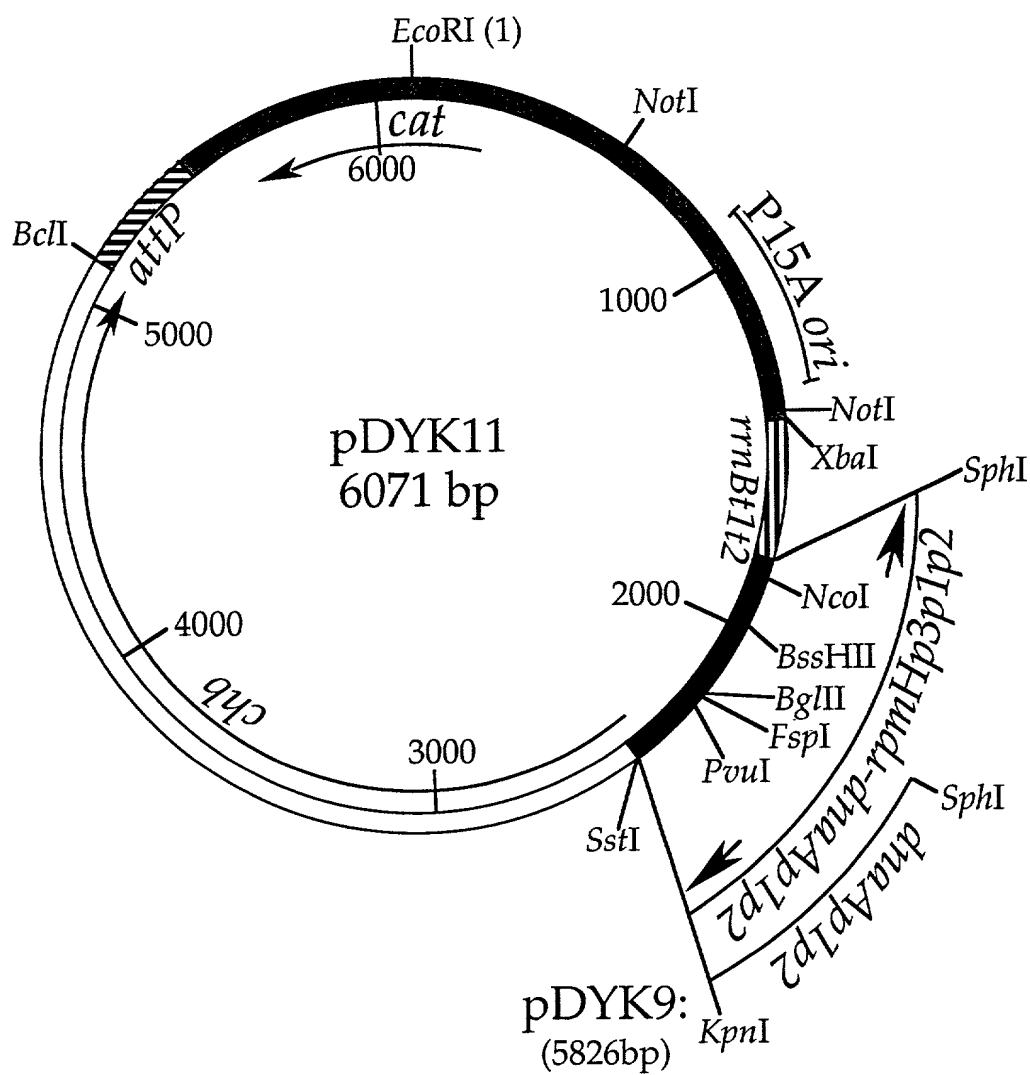
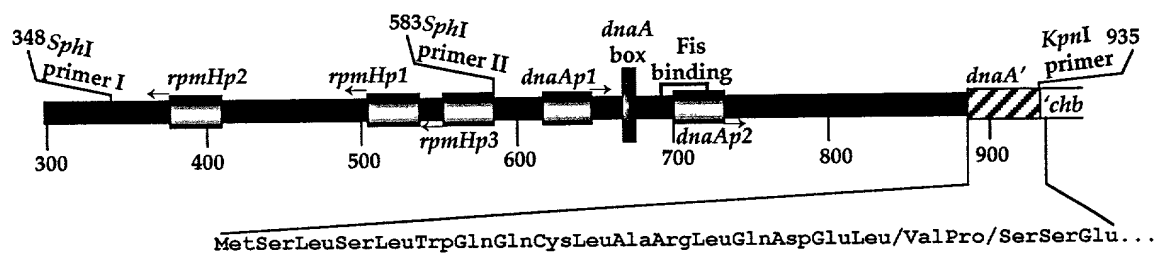


Figure 3



**Figure 4**



SEQ ID NO: 17

Figure 5

# SEQUENCE LISTING

<110> Zyskind, Judith W.

<120> CHITOBIASE AS A REPORTER ENZYME

<130> ELITRA.006A

<150> 60/159,221

<151> 1999-10-13

<160> 19

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 1

gcacatgcat gctgggtcatt aaattttcc

29

<210> 2

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 2

cggggtacca actcatcctg caatcg

26

<210> 3

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 3

ctagtctaga tgccgaactc agaagtga

28

<210> 4

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 4	
gcacatgcat gcgggggatg gcttgtagat	30
<210> 5	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Oligonucleotide	
<400> 5	
catgatcatg cgacagggtt gatga	25
<210> 6	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Oligonucleotide	
<400> 6	
gggggcgcct acctttcacg ag	22
<210> 7	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Oligonucleotide	
<400> 7	
catgcatgca tgaaacgatg gacacc	26
<210> 8	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Oligonucleotide	
<400> 8	
cggggtacca actcatcctg caatcg	26
<210> 9	
<211> 12	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Oligonucleotide	
<400> 9	
cattaatgca tg	12



<210> 10  
 <211> 12  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Oligonucleotide

<400> 10  
 cattaatgca tg

12

<210> 11  
 <211> 5826  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> pDYK9

<400> 11  
 gaattccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg gataaaaactt 60  
 gtgcttattt ttctttacgg tcttttaaaa ggccgtaata tccagctgaa cggctctggtt 120  
 ataggtacat tgagcaactg actgaaatgc ctcaaaatgt tctttacgat gccattggga 180  
 tatatcaacg gtggtatata cagtgatttt tttctccatt ttagcttcct tagctcctga 240  
 aaatctcgat aactcaaaaa atacgcccgg tagtgatctt atttcattat ggtgaaagtt 300  
 ggaacctctt acgtgccgat caacgtctca ttttcgccaa aagttggccc agggcttccc 360  
 ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggat 420  
 ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt 480  
 gtttttgagg tgctccagtg gcttctggtt ctatcagctg tccctcctgt tcagctactg 540  
 acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtgcggc 600  
 cgactggct tactatggtg gcactgatga ggggtgcagt gaagtgcctc atgtggcagg 660  
 agaaaaaagg ctgcaccggg gcgtcagcag aatatgtgat acaggatata ttccgcttcc 720  
 tcgctcactg actcgtacg ctcggtcggt cgactgcggc gagcggaaat ggcttacgaa 780  
 cggggcggag atttcctgga agatgccagg aagatactta acagggaagt gagaggccg 840  
 cggcaaagcc gtttttccat aggtccggcc cccctgacaa gcatcacgaa atctgacgct 900  
 caaatcagtg gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggcgg 960  
 ctccctcggt cgctctcctg ttctgcctt tcggtttacc ggtgtcattc cgctggtatg 1020  
 gccgcgtttg tctcattcca cgcctgacac tcagttccgg gtaggcagtt cgctccaagc 1080  
 tggactgtat gcacgaaccc cccgttcagt ccgaccgctg cgcttattcc ggtaactatc 1140  
 gtcttgagtc caaccgggaa agacatgcaa aagcaccact ggcagcagcc actggttaatt 1200  
 gatttagagg agttagtctt gaagtcatgc gccgggttaag gctaaactga aaggacaagt 1260  
 tttggtgact gcgctcctcc aagccagtta cctcggttca aagagttggt agctcagaga 1320  
 accttcgaaa aaccgccctg caaggcgggt ttttcgtttt cagagcaaga gattacgcgc 1380  
 agaccaaacc gatctcaaga agatcatctt atgcggccgc atcagataaa atatttctag 1440  
 atgccgaact cagaagtga acgcccgtag gccgatggta gtgtggggtc tccccatgcg 1500  
 agagtaggga actgccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt 1560  
 tcgttttata tgttgtttgt cgggtgaacgc tctcctgagt aggacaaatc cgccgggagc 1620  
 ggatttgaac gttgcgaagc aacggcccgg aggggtggcg gcaggacgcc cgccataaac 1680  
 tgccaggcat caaattaagc agaaggccat cctgacggat ggcttttttg cgtttctaca 1740  
 aactcttctt gtcgtcatat ctacaagcca tccccgcgat gctggtcatt aaattttcca 1800  
 atatgcggcg taaatcgtgc ccgcctcgcg gcaggatcgt ttacacttag cgagttctgg 1860  
 aaagtccctg ggataaatcg ggaaaatctg tgagaaacag aagatctctt gcgcagttta 1920  
 ggctatgata cgcggtcccc atcgttttgc aggatcttga tcgggcatat aaccgcagac 1980  
 agcggttcgt gcgtcaccct caagcagggt cttttcgacg tacgtcaaca atcatgaatg 2040  
 tttcagcctt agtcattatc gacttttgtt cgagtggagt ccgcggtgtc actttcgcctt 2100  
 tggcagcagt gtcttgcccc attgcaggat gagttggtac cgagctctga gcaacaagtt 2160

gtaaactcac	tggctgataa	ccttgatata	caatatgaag	tgtaaactaa	ccatggtgct	2220
aacgaaggtc	ttgcgtgcca	agatatgggc	gcagaatggg	cttcttgtaa	caaagtaaac	2280
atgacgcttg	ttaaccaagg	tgaagctggt	gactcaaaaag	attgggctat	ttacttccac	2340
agcattcgtc	tgattctgga	tggtgacaac	gagcagttca	aaatctctcg	tgtaacgggt	2400
gacctacata	agctagaacc	aacagataag	tttgacggct	tcgctgccgg	tgaagagggt	2460
gttcttccat	tggttggtga	atactggcaa	ctatttgaaa	ctgacttcat	gccgggtgca	2520
ttcgtttctg	ctccaaacgc	agaacctaag	atgattgctt	ctctaaatac	tgaagatggt	2580
gcgtcttttg	tgacgggtct	tgaaggtaac	aacctaaaaac	gtacaccaga	tgacaacaat	2640
gtatttgcaa	acgctgtgtc	tcgttttgag	aaaaacgaag	acctagcaac	acaagacgta	2700
tcaaccacgt	tactaccaac	accaatgcac	ggtgaagcgg	gtaaaggcaa	agtagatata	2760
goggatggta	ttgcgctgcc	taaagacgca	ttcgatgcga	ctcagttcgc	agcgattcaa	2820
gatcgtgcag	aagtggtagg	tgtggacggt	cgtgggtgatc	ttcctgtaag	catcactggt	2880
gttctcgcag	acttcaccgg	tgaattagca	aaatctggtg	cttacgaaat	gagcatcaaa	2940
ggcgacggta	ttgtgattaa	agcgttcgac	caagcaggcg	ctttctacgc	agtacaatct	3000
atctttggcc	tggtagatag	ccaaaatgct	gattctctac	cacaactgtc	tattaaagat	3060
gcgcctcggt	ttgattaccg	tggtgtgatg	gtggatgtgg	ctcgtaactt	ccactctaag	3120
gacgcaatcc	ttgcaacgct	agaccaaagt	gcagcgtaca	agatgaacaa	acttcacctt	3180
cacctaacgg	atgatgaagg	ctggcggtta	gaaatcccgg	gtctgcctga	gctgacagaa	3240
gtgggtgcta	accgttggtt	cgatacacaa	gagaaaagct	gtttactgcc	tcagcttggc	3300
tctggtccaa	cgacagacaa	ctttggctct	ggctacttca	gcaaagcaga	ctacgtggaa	3360
atcttgaaat	acgcgaaagc	acgtaacatt	gaagtgtattc	cagaaatcga	tatgccagct	3420
cacgctcgtg	cagcagtagt	atcaattgga	gctcgttacg	accgcctaata	ggaagaagggt	3480
aaagaagctg	aagcgaacga	ataccgtctg	atggatcctc	aagatacatc	aaacgtaacg	3540
acggttcagt	tctacaataa	gcaaagcttc	atcaacccat	gtatggaatc	ttcaactcgc	3600
tttgttgata	aggtgatttc	agaagtggca	gcaatgcacc	aagaagctgg	cgctccacta	3660
acaacttggc	acttcggtgg	tgacgaagcg	aagaacatca	agctaggtgc	tggtttccaa	3720
gacgttaacg	cagaagataa	agtaagctgg	aaaggcacga	ttgacctgtc	taaacaagac	3780
aagccgtttg	cacagtctcc	acaatgtcag	acgctaatac	cagatggcac	agtcagtgc	3840
tttgctcacc	taccaagcca	cttcgcggaa	gaagtgtcga	agattgttgc	tgagaaaggc	3900
attccaaact	tccaagcttg	gcaagatggt	ttgaaataca	gtgacggcga	aaaagcgttc	3960
gctacagaaa	atactcgcgt	aaacttctgg	gacgttctgt	actggggcgg	tacttcctca	4020
gtgtacgagt	ggtctaagaa	aggttacgac	gtgattgttt	ctaaccacga	ttacgtgtac	4080
atggatatgc	catacgaagt	tgacccgaaa	gagcgtgggt	actactgggc	aacacgtgca	4140
acggatactc	gtaagatggt	tggttttgca	ccagagaaca	tgccctcaaaa	cgagaaact	4200
tctgtagatc	gcgatggcaa	tggttttact	ggtaaagggtg	aaatcgaagc	gaaacctttc	4260
tacggtctat	ctgcacaact	ttggtctgag	acagtaagta	acgacgagca	atacgagtac	4320
atggtatttc	ctcgcgtcct	cgctcgtgct	cagcgtgcat	ggcaccgtgc	tgactgggaa	4380
aacgactaca	aagttgggtg	tgagtactcg	caaaactcta	atctagtgtg	taaagcatcg	4440
ctaaaccaag	actacaacgg	ctttgcgaac	gtacttggtc	aacgtgaact	ggctaagcta	4500
gaaaaatcag	gtattgacta	ccgcctacca	gtaccagggtg	caaaagtaga	agatggtaag	4560
ctagcaatga	acgttcagtt	ccctggcgta	acgcttcaat	actctctgga	tggtgagaac	4620
tggttgactt	atgcagacaa	cgctcgtcca	aatgttaactg	gtgaagtctt	catccgctcg	4680
gtatctgcga	caggtgagaa	ggtaagccgt	atcactagcg	tgaaataata	gcgctcagta	4740
ttcactaaaa	tcatagttcc	ttactcaaag	ccctcaactt	atggtggggg	ctttgtttat	4800
ttttcttcgg	aaaataagcg	tgatcatgcg	acaggtttga	tgacaaaaaa	ttagcgcaag	4860
aagacaaaaa	tcaccttgcg	ctaattgctct	gttacagggtc	actaatacca	tctaagtagt	4920
tgattcatag	tgactgcata	tggtgtgttt	tacagtatta	tgtagtctgt	tttttatgca	4980
aaatctaatt	taatataattg	atatattatat	cattttacgt	ttctcgttca	gcttttttat	5040
actaagttgg	cattataaaaa	aagcattgct	tatcaatttg	ttgcaacgaa	caggtcacta	5100
tcagtcaaaa	taaaatcatt	atttgatttc	aattttgtcc	cactccctgc	ctctgtcatc	5160
acgatactgt	gatgccatgg	tgtccgactt	atgcccgaga	agatggttag	caaacttatc	5220
gcttatctgc	ttctcataga	gtcttgacga	caaactgcgc	aactcgtgaa	aggtaggcgc	5280
ccccggtcga	atttgctttc	gaatttctgc	cattcatccg	ottattatca	cttattcagg	5340
cgtagcacca	ggcgtttaag	ggcaccaata	actgccttaa	aaaaattacg	ccccgcctg	5400
ccactcatcg	cagtactggt	gtaattcatt	aagcattctg	ccgacatgga	agccatcaca	5460
gacggcatga	tgaacctgaa	tcgccagcgg	catcagcacc	ttgtcgcctt	gcgtataata	5520
tttgcccatg	gtgaaaacgg	ggcggaagaa	gttgtccata	ttggccacgt	ttaaatcaaa	5580

actggtgaaa	ctcaccag	gattggctga	gacgaaaaac	atattctcaa	taaacccttt	5640
agggaaatag	gccagggtttt	caccgtaaca	cgccacatct	tgcgaaatata	tgtgtagaaa	5700
ctgccggaaa	tcgtcgtggt	attcactcca	gagcgatgaa	aacgtttcag	tttgctcatg	5760
gaaaacggtg	taacaagggt	gaacactatc	ccatatacacc	agctcaccgt	ctttcattgc	5820
catacg						5826

<210> 12  
 <211> 6071  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> pDYK11

<400> 12	
gaattccgga	tgagcattca
gtgcttattt	ttcttttacg
ataggtagat	tgagcaactg
tatatcaacg	gtggtatata
aaatctcgat	aactcaaaaa
ggaacctctt	acgtgcccag
ggtatcaaca	gggacaccag
ttattcggcg	caaagtgcgt
gtttttgagg	tgctccagtg
acggggtggt	gcgtaacggc
cgcactggct	tactatgttg
agaaaaaagg	ctgcaccggt
tcgctcactg	actcgctacg
cggggcggag	atttctctga
cggcaaagcc	gtttttccat
caaatcagtg	gtggcgaaac
ctccctcggt	cgctctcctg
gccgcgtttg	tctcattcca
tggtactgtat	gcacgaaccc
gtcttgagtc	caacccggaa
gatttagagg	agttagtcct
tttggtgact	gcgctcctcc
accttcgaaa	aaccgcccgt
agacaaaaac	gatctcaaga
atgccgaact	cagaagtga
agagtaggga	actgccaggc
tcgtttttat	tggtgtttgt
ggatttgaac	gttgcgaaag
tgccaggcat	caaattaagc
aactcttcct	gtcgtcata
aagccatggg	tgattaaa
tttccttatt	taccgcgctt
agtggaaacc	agcatggcgc
ggatcgctct	ggatctttat
ttccaatatg	cggcgtaaat
tctggaaagt	cctgtggata
gtttaggcta	tgatccgcgg
cagacagcgg	ttcgtgcgtc
gaatgtttca	gccttagtca
cgctttggca	gcagtgtctt
aagttgtaaa	ctcactggct
gtgctaacga	aggtcttgcg
tcaggcgggc	aagaatgtga
ggccgtaata	tccagctgaa
ctcaaaaatgt	tcctttacgat
ttctctccatt	ttagcttcct
tagtgatctt	atttcattat
ttttcgccaa	aagttggccc
ttctgcgaag	tgatcttcog
tgccaactta	ctgatttagt
ctatcagctg	tcctcctgt
cgggacatca	gcgctagcgg
gaagtgcctc	atgtggcagg
aatatgtgat	acaggatata
cgactgcggc	gagcggaaat
aagatactta	acagggaagt
cccctgacaa	gcatcacgaa
tataaagata	ccaggcgctt
tcggtttacc	ggtgtcattc
tcagttccgg	gtaggcagtt
cgccttcagt	ccgaccgctg
agacatgcaa	aagcaccact
gccgggttaag	gctaactga
cctcggttca	aagagttggt
ttttcgcttt	cagagcaaga
atgcggccgc	atcagataaa
gocgatggta	gtgtggggtc
tcctctgagt	aggacaaatc
aggggtggcg	gcaggacgcc
cctgacggat	ggcctttttg
tcacaaagcca	gcataaaac
taataattgt	acactccgga
tttcgcaggg	aaaatgtacg
ggattatacg	ggctgatggg
aagccaattt	ttgtctatgg
tcgcggcagg	atcgtttaca
aatcggaata	atctgtgaga
tttgaggat	cttgcaggat
aggggtcttt	cgacgtacgt
ggagtcggcc	gtgtcacttt
aggatgagtt	ggtaccgagc
atatccaata	tgaagtgtta
tgaggcgcaga	atgggcttct
	tgtaacaaag

taaacatgac	gcttggttaac	caaggtgaag	ctggttgactc	aaaagattgg	gctattttact	2580
tccacagcat	tcgtctgatt	ctggatggtg	acaacgagca	gttcaaaatc	tctcgtgtaa	2640
cgggtgacct	acataagcta	gaaccaacag	ataagtttga	cggcttcgct	gccggtgaag	2700
agggtgttct	tccattgggt	ggtgaatact	ggcaactatt	tgaaactgac	ttcatgccgg	2760
gtgcattcgt	ttctgctcca	aacgcagaac	ctaagatgat	tgcttctcta	aatactgaag	2820
atggtgcgtc	ttttgtgacg	ggtcttgaag	gtaacaacct	aaaacgtaca	ccagatgaca	2880
acaatgtatt	tgcaaacgct	gtgtctcggt	ttgagaaaaa	cgaagacct	gcaacacaag	2940
acgtatcaac	cacgttacta	ccaacaccaa	tgacacgttg	agcgggtaaa	ggcaaagtag	3000
atatcgcgga	tggtattgcy	ctgcctaaag	acgcattoga	tgcgactcag	ttcgcgagca	3060
ttcaagatcg	tgacagaagt	gtagggtgtg	acgttcggtg	tgatcttcct	gtaagcatca	3120
ctggtgttcc	tgacagacttc	accggtgaat	tagcaaaatc	tggtgcttac	gaaatgagca	3180
tcaaaggcga	cggatattgt	attaaagcgt	tcgaccaagc	aggcgctttc	tacgcagtac	3240
aatctatctt	tggcctggta	gatagccaaa	atgctgattc	tctaccacaa	ctgtctatta	3300
aagatgcgcc	tcgttttgat	taccgtgggt	tgatgggtga	tgtggctcgt	aacttccact	3360
ctaaggacgc	aatccttgca	acgctagacc	aaatggcagc	gtacaagatg	aacaaacttc	3420
accttcacct	aaccgatgat	gaaggctggc	gtttagaaat	cccgggtctg	cctgagctga	3480
cagaagtggg	tgctaaccgt	tgtttcgata	cacaagagaa	aagctgttta	ctgcctcagc	3540
ttggctctgg	tccaacgaca	gacaactttg	gctctggcta	cttcagcaaa	gcagactacg	3600
tggaaatctt	gaaatacgcg	aaagcacgta	acattgaagt	gattccagaa	atcgatatgc	3660
cagctcacgc	tcgtgcagca	gtagtatcaa	tggaaagctc	ttacgaccgc	ctaattggaag	3720
aaggtaaaga	agctgaagcg	aacgaatacc	gtctgatgga	tcctcaagat	acatcaaacg	3780
taacgacggg	tcagttctac	aataagcaaa	gcttcatcaa	cccatgtatg	gaatcttcaa	3840
ctcgctttgt	tgataagggt	atttcagaag	tggcagcaat	gcaccaagaa	gctggcgctc	3900
cactaacaac	ttggcacttc	ggtggtgacg	aagcgaagaa	catcaagcta	ggtgctggtt	3960
tccaagacgt	taacgcagaa	gataaagtaa	gctggaaagg	cacgattgac	ctgtctaaac	4020
aagacaagcc	gtttgcacag	tctccacaat	gtcagacgct	aatcacagat	ggcacagtca	4080
gtgactttgc	tcacctacca	agccacttcg	cggaaagaag	gtcgaagatt	gttgctgaga	4140
aaggcattcc	aaacttccaa	gcttggcaag	atggtttgaa	atacagtacg	ggcgaaaaag	4200
cgttcgctac	agaaaatact	cgcgtaaact	tctgggacgt	tctgtactgg	ggcggtactt	4260
cctcagtgtg	cgagtgggtc	aagaaagggt	acgacgtgat	tgtttctaac	ccagattacg	4320
tgtacatgga	tatgccatac	gaagttgacc	cgaagagcgc	tggttactac	tgggcaacac	4380
gtgcaacgga	tactcgtaag	atgtttggct	ttgcaccaga	gaacatgcct	caaaacgcag	4440
aaacttctgt	agatcgcgat	ggcaatggct	ttactggtaa	aggtgaaatc	gaagcgaaac	4500
ctttctacgg	tctatctgca	caactttggt	ctgagacagt	acgtaacgac	gagcaatacg	4560
agtacatggt	attccctcgc	gtcctcgctg	ctgctcagcg	tgcatggcac	cgtgctgact	4620
gggaaaacga	ctacaaagtt	ggtgttgagt	actcgcaaaa	ctctaattcta	gttgataaag	4680
catcgctaaa	ccaagactat	aaccgctttg	cgaacgtact	tggtcaacgt	gaactggcta	4740
agctagaaaa	atcaggtatt	gactaccgcc	taccagtacc	aggtgcaaaa	gtagaagatg	4800
gtaagctagc	aatgaacgtt	cagttccctg	gcgtaacgct	tcaataactct	ctggatgggt	4860
agaactggtt	gacttatgca	gacaacgctc	gtccaaatgt	aactggtgaa	gtcttcatcc	4920
gctcggtatc	tgcgacaggt	gagaaggtaa	gccgtatcac	tagcgtgaaa	taatagcgct	4980
cagtattcac	taaaatcata	gttccttact	caaagccctc	aacttatggt	gggggctttg	5040
tttatttttc	ttcggaaaat	aagcgtgatc	atgcgacagg	tttgatgaca	aaaaatttagc	5100
gcaagaagac	aaaaatcacc	ttgcgctaata	gctctgttac	aggtcactaa	taccatctaa	5160
gtagttgatt	catagtgact	gcatatgttg	tgttttacag	tattatgtag	tctgtttttt	5220
atgcaaaatc	taattttaata	tattgatatt	tatatcattt	tacgtttctc	gttcagcttt	5280
tttatactaa	gttggcatta	taaaaaagca	ttgcttatca	atttggtgca	acgaacaggt	5340
cactatcagt	caaaataaaa	tcattatttg	atttcaattt	tgtoocactc	cctgcctctg	5400
tcacacagat	actgtgatgc	catggtgtcc	gacttatgcc	cgagaagatg	ttgagcaaac	5460
ttatcgctta	tctgcttctc	atagagtctt	gcagacaaac	tgcgcaactc	gtgaaaggta	5520
ggcgcccccg	gtccgaattt	ctttcgaatt	tctgccattc	atccgcttat	tatcacttat	5580
tacggcgtag	caccaggcgt	ttaagggcac	caataactgc	cttaaaaaaa	ttacgccccg	5640
ccctgccact	catcgcgagta	ctgttgtaat	tcattaagca	ttctgccgac	atggaagcca	5700
tcacagacgg	catgatgaac	ctgaatcgcc	agcggcatca	gcacottgtc	gccttgcgta	5760
taatatttgc	ccatggtgaa	aacgggggcg	aagaagttgt	ccatattggc	cacgtttaaa	5820
tcaaaactgg	tgaaactcac	ccagggattg	gctgagacga	aaaacatatt	ctcaataaac	5880
cctttaggga	aataggccag	gttttcaccg	taacacgcca	catcttgcca	atatatgtgt	5940

agaaactgcc	ggaaatcgtc	gtggtattca	ctccagagcg	atgaaaacgt	ttcagtttgc	6000
tcatggaaaa	cggtgtaaca	agggtgaaca	ctatcccata	tcaccagctc	accgtctttc	6060
attgccatac	g					6071
<210> 13						
<211> 5641						
<212> DNA						
<213> Artificial Sequence						
<220>						
<223> pJMF3						
<400> 13						
gaattccgga	tgagcattca	tcaggcgggc	aagaatgtga	ataaaggccg	gataaaaactt	60
gtgcttattt	ttcttttacg	tctttaaaaa	ggccgtaata	tccagctgaa	cggtctgggt	120
ataggtacat	tgagcaactg	actgaaatgc	ctcaaaatgt	tctttacgat	gccattggga	180
tatatcaacg	gtggtataat	cagtgtattt	tttctccatt	ttagcttctt	tagctcctga	240
aaatctcgat	aactcaaaaa	atacgcccgg	tagtgatctt	atttcattat	ggtgaaagtt	300
ggaacctctt	acgtgccgat	caacgtctca	ttttcgccaa	aagttggccc	agggcttccc	360
ggtatcaaca	gggacaccag	gatttattta	ttctgcgaag	tgatcttccg	tcacagggtat	420
ttattcggcg	caaagtgcgt	cgggtgatgc	tgccaaactta	ctgatttagt	gtatgatggt	480
gtttttgagg	tgctccagtg	gcttctgttt	ctatcagctg	tccctcctgt	tcagctactg	540
acggggtggt	gcgtaacggc	aaaagcaccg	ccggacatca	gcgctagcgg	agtgtgcggc	600
cgcactggct	tactatgttg	gcactgatga	gggtgtcagt	gaagtgcctc	atgtggcagg	660
agaaaaaagg	ctgcaccggt	gcgtcagcag	aatatgtgat	acaggatata	ttccgcttcc	720
tcgctcactg	actcgctacg	ctcggtcggt	cgaactgcgg	gagcggaaat	ggcttacgaa	780
cggggcggag	atttcctgga	agatgccagg	aagatactta	acagggaagt	gagagggccg	840
cggcaaagcc	gtttttccat	aggtcccgcc	cccctgacaa	gcatacagaa	atctgacgct	900
caaatcagtg	gtggcgaaac	ccgacaggac	tataaagata	ccaggcggtt	cccctggcgg	960
ctccctcgtg	cgtctcctg	ttcctgcctt	tcggtttacc	ggtgtcattc	cgctgttatg	1020
gccgcgtttg	tctcattcca	cgcctgacac	tcagttccgg	gtaggcagtt	cgctccaagc	1080
tggactgtat	gcacgaaccc	cccgttcagt	ccgaccgctg	cgccttatcc	ggtaactatc	1140
gtcttgagtc	caacccggaa	agacatgcaa	aagcaccact	ggcagcagcc	actggttaatt	1200
gatttagagg	agtttagctt	gaagtcatgc	gccggttaag	gctaaactga	aaggacaagt	1260
tttggtgact	gcgtcctctc	aagccagtta	cctcggttca	aagagttggt	agctcagaga	1320
accttcgaaa	aaccgccttg	caaggcggtt	ttttcgtttt	cagagcaaga	gattacgcgc	1380
agaccaaacc	gatctcaaga	agatcatctt	atgcggccgc	atcagataaa	atattttctag	1440
atgccgaact	cagaagtga	acgccgtagc	gcgatggta	gtgtggggtc	tccccatgcg	1500
agagttagga	actgcaggcg	atcaaataaa	acgaaggcgt	cagtcgaaag	actgggcctt	1560
tcgttttatc	tgttgtttgt	cgttgaaccc	tgctcctgag	aggacaaatc	cgccggggagc	1620
ggatttgaac	gttgcggaag	aacggcccgc	atgggtggcg	gcaggacgcc	cgccataaac	1680
tgccaggcat	caaattaagc	agaaggccat	cctgacggat	ggcctttttg	cgtttctaca	1740
aactcttctt	gtcgtcatat	ctacaagcca	tccccgcgat	gcattaatgt	gagttagctc	1800
actcattagg	caccccaggc	tttacacttt	atgcttccgg	ctcgtatggt	gtgtggaatt	1860
gtgagcggat	aacaatttca	cacaggaaac	agctatgacc	atgattacgc	caagcttgca	1920
tgcctgcagg	tcgactctag	aggatccccg	ggtaccgagc	tctgagcaac	aagttgtaaa	1980
ctcactggct	gataaccttg	atatccaata	tgaagtgtta	actaaccatg	gtgctaacga	2040
aggtcttgcg	tgccaagata	tgggcgcaga	atgggcttct	tgtaacaaag	taaacatgac	2100
gcttggttaac	caaggtgaag	ctggtgactc	aaaagattgg	gctattttact	tccacagcat	2160
tcgtctgatt	ctggatgttg	acaacgagca	gttcaaaatc	tctcgtgtta	cgggtgacct	2220
acataagcta	gaaccaacag	ataagtttga	cggcttcgct	gccggtgaag	aggttggtct	2280
tccattgggt	ggtgaatact	ggcaactatt	tgaactctgc	ttcatgccgg	gtgcattcgt	2340
ttctgtccca	aacgcagaac	ctaagatgat	tgcttctcta	aatactgaag	atgttgcgct	2400
ttttgtgacg	ggtcttgaag	gtaacaacct	aaaacgtaca	ccagatgaca	acaatgtatt	2460
tgcaaacgct	gtgtctcggt	ttgagaaaaa	cgaagacctc	gcaacacaag	acgtatcaac	2520
cacgttacta	ccaacaccaa	tgacggttga	agcgggtaaa	ggcaaagtag	atatcgcgga	2580
tggtattgcg	ctgcctaaag	acgcattcga	tgcgactcag	ttcgcagcga	ttcaagatcg	2640
tgcagaagtg	gtaggtgtgg	acgttcgtgg	tgatcttctt	gtaagcatca	ctgttggttc	2700

tgcagacttc	accggtgaat	tagcaaaatc	tggtgcttac	gaaatgagca	tcaaaggcga	2760
cggatttg	attaaagcgt	tcgaccaagc	aggcgctttc	tacgcagtac	aatctatctt	2820
tggcctggta	gatagccaaa	atgctgattc	tctaccacaa	ctgtctatta	aagatgcgcc	2880
tcgttttgat	taccgtgggtg	tgatgggtga	tgtggctcgt	aacttccact	ctaaggacgc	2940
aatccttgca	acgctagacc	aaatggcagc	gtacaagatg	aacaaacttc	accttcacct	3000
aaccgatgat	gaaggctggc	gtttagaaat	cccggtctg	cctgagctga	cagaagtggg	3060
tgtaaccgt	tgtttcgata	cacaagagaa	aagctgttta	ctgcctcagc	ttggctctgg	3120
tccaacgaca	gacaactttg	gctctggcta	cttcagcaaa	gcagactacg	tggaatctt	3180
gaaatacgcg	aaagcacgta	acattgaagt	gattccagaa	atcgatatgc	cagctcacgc	3240
tcgtgcagca	gtagtatcaa	tggagctcg	ttacgaccgc	ctaattggaag	aaggtaaaga	3300
agctgaagcg	aacgaatacc	gtctgatgga	tcctcaagat	acatcaaacg	taacgacggg	3360
tcagttctac	aataagcaaa	gcttcatcaa	cccatgtatg	gaatcttcaa	ctcgctttgt	3420
tgataaggtg	atttcagaag	tggcagcaat	gcaccaagaa	gctggcgctc	cactaacaac	3480
ttggcacttc	ggtggtgacg	aagcgaagaa	catcaagcta	ggtgctgggt	tccaagacgt	3540
taacgcagaa	gataaagtaa	gctggaaagg	cacgattgac	ctgtctaaac	aagacaagcc	3600
gtttgcacag	tctccacaat	gtcagacgct	aatcacagat	ggcacagtca	gtgactttgc	3660
tcacctacca	agccacttcg	cggaagaagt	gtcgaagatt	gttgctgaga	aaggcattcc	3720
aaacttccaa	gcttggcaag	atggtttgaa	atacagtgac	ggcgaaaaag	cgttcgctac	3780
agaaaatact	cgcgtaaaact	tctgggacgt	tctgtactgg	ggcggtactt	cctcagtgtg	3840
cgagtggctc	aagaaaaggt	acgacgtgat	tgtttctaac	ccagattacg	tgtacatgga	3900
tatgccatac	gaagttgacc	cgaaagagcg	tggttactac	tgggcaacac	gtgcaacgga	3960
tactcgtaag	atgtttggct	ttgcaccaga	gaacatgcct	caaaacgcag	aaacttctgt	4020
agatcgcgat	ggcaatggct	ttactggtaa	aggtgaaatc	gaagcgaaac	ctttctacgg	4080
tctatctgca	caactttggt	ctgagacagt	acgtaacgac	gagcaatacg	agtacatggg	4140
attccctcgc	gtcctcgtcg	ctgctcagcg	tgcattggac	cgtgctgact	gggaaaacga	4200
ctacaaagtt	ggtgttgagt	actcgcaaaa	ctctaatact	gttgataaag	catcgctaaa	4260
ccaagactac	aaccgctttg	cgaacgtact	tggccaacgt	gaactggcta	agctagaaaa	4320
atcaggtatt	gactaccgcc	taccagtacc	aggtgcaaaa	gtagaagatg	gtaagctagc	4380
aatgaacggt	cagttccctg	gcgtaacgct	tcaatactct	ctggatgggt	agaactgggt	4440
gacttatgca	gacaacgctc	gtccaaatgt	aactggtgaa	gtcttcatcc	gctcgggtac	4500
tgcgacaggt	gagaaggtaa	gccgtatcac	tagcgtgaaa	taatagcgct	cagtattcac	4560
taaaatcata	gttccttact	caaagccctc	aacttatgtt	gggggctttg	tttatttttc	4620
ttcgaaaaat	aagcgtgatc	atgcgacagg	tttgatgaca	aaaaattagc	gcaagaagac	4680
aaaaatcacc	ttgcgctaatt	gctctgttac	aggtcactaa	taccatctaa	gtagttgatt	4740
catagtgaat	gcataatgtt	tggtttacag	tattatgtag	tctgtttttt	atgcaaaatc	4800
taatttaata	tattgatatt	tatatcattt	tacgtttctc	gttcagcttt	tttatactaa	4860
gttggcatta	taaaaaagca	ttgcttatca	atttgttgca	acgaacaggt	cactatcagt	4920
caaaaataaaa	tcattattttg	atttcaattt	tgtcccactc	cctgcctctg	tcatacagat	4980
actgtgatgc	catggtgtcc	gacttatgcc	cgagaagatg	ttgagcaaac	ttatcgctta	5040
tctgcttctc	atagagtctt	gcagacaaac	tgcgcaactc	gtgaaaggta	ggcgcccccg	5100
gtcgaatttg	ctttcgaatt	tctgccattc	atccgcttat	tatcaattat	tcaggcgtag	5160
caccaggcgt	ttaagggcac	caataactgc	cttaaaaaaa	ttacgccccg	ccctgccact	5220
catcgagta	ctgttgtaat	tcattaagca	ttctgccgac	atggaagcca	tcacagacgg	5280
catgatgaac	ctgaatcgcc	agcggcatca	gcaccttgct	gccttgcgta	taatatttgc	5340
ccatggtgaa	aacgggggcg	aagaagtgt	ccatattggc	cacgtttaaa	tcaaaactgg	5400
tgaaactcac	ccagggtattg	gctgagacga	aaaacatatt	ctcaataaac	cctttaggga	5460
aataggccag	gttttcaccg	taacacgcca	catcttgcca	atatatgtgt	agaaactgcc	5520
ggaaatcgct	gtggtattca	ctccagagcg	atgaaaacgt	ttcagtttgc	tcattgaaaa	5580
cggtgtaaca	aggggtgaaca	ctatcccata	tcaccagctc	accgtctttc	attgccatac	5640
g						5641

<210> 14

<211> 5670

<212> DNA

<213> Artificial Sequence

<220>

<223> pJMF4

<400> 14

gaattccgga	tgagcattca	tcaggcgggc	aagaatgtga	ataaaggccg	gataaaaactt	60
gtgcttattt	ttctttacgg	tcttttaaaa	ggcgcgaata	tccagctgaa	cggtctgggtt	120
ataggtacat	tgagcaactg	actgaaatgc	ctcaaaatgt	tctttacgat	gccattggga	180
tatatcaacg	gtggtatatc	cagtgatctt	tttctccatt	ttagcttcct	tagctcctga	240
aaatctcgat	aactcaaaaa	atacgcgcgg	tagtgatctt	atttcattat	ggtgaaagtt	300
ggaacctctt	acgtgccgat	caacgtctca	ttttcgccaa	aagttggccc	agggcttccc	360
ggtatcaaca	gggacaccag	gattttattta	ttctgcgaag	tgatcttccg	tcacaggtat	420
ttattcggcg	caaagtgcgt	cggtgatgc	tgccaaactta	ctgatttagt	gtatgatggt	480
gtttttgagg	tgctccagtg	gcttctgttt	ctatcagctg	tccctcctgt	tcagctactg	540
acggggtggg	gcgtaacggc	aaaagcaccg	ccggacatca	gcgctagcgg	agtgtgcggc	600
cgcactggct	tactatgttg	gcactgatga	gggtgtcagt	gaagtgcctc	atgtggcagg	660
agaaaaaagg	ctgcaccggg	gcgtcagcag	aatatgtgat	acaggatata	ttccgcttcc	720
tcgctcactg	actcgctacg	ctcggtcggt	cgactgcggc	gagcggaaat	ggcttacgaa	780
cggggcggag	atttctctga	agatgccagg	aagatactta	acaggggaagt	gagagggccg	840
cggcaaagcc	gtttttccat	aggctccgcc	cccctgacaa	gcatacacgaa	atctgacgct	900
caaatacagt	gtggcgaaac	ccgacaggac	tataaagata	ccaggcggtt	cccctggcgg	960
ctccctcgtg	cgctctcctg	ttcctgcctt	togggtttacc	ggtgtcattc	cgctgttatg	1020
gccgcgtttg	tctcattcca	cgcctgacac	tcagttccgg	gtaggcagtt	cgctccaagc	1080
tgactgtat	gcacgaaccc	cccgttcagt	ccgaccgctg	cgcttatcc	ggttaactatc	1140
gtcttgagtc	caacccgga	agacatgcaa	aagcaccact	ggcagcagcc	actggttaatt	1200
gttttagagg	agttagctt	gaagtcatgc	cgcggttaag	gctaaactga	aaggacaagt	1260
tttggtgact	gcgctcctcc	aagccagtta	cctcggttca	aagagttggt	agctcagaga	1320
accttcgaaa	aaccgccctg	caaggcgggt	ttttcgtttt	cagagcaaga	gattacgcgc	1380
agacaaaac	gatctcaaga	agatcatctt	atgcggccgc	atcagataaa	atatttctag	1440
atgccgaact	cagaagtga	acgccgtagc	gccgatggta	gtgtggggtc	tccccatgcg	1500
agagtaggga	actgccaggc	atcaaataaa	acgaaaggct	cagtcgaaag	actgggcctt	1560
tcgttttatc	tgttgtttgt	cgtgaaacgc	tctcctgagt	aggacaaatc	cgccggggagc	1620
ggatttgaac	gttgcggaagc	aacggcccgg	aggggtggcg	gcaggacgcc	cgccataaac	1680
tgccaggcat	caaattaagc	agaaggccat	cctgacggat	ggcctttttg	cgtttctaca	1740
aactcttcct	gtcgtcatat	ctacaagcca	tccccgcgat	gcattaatgt	gagttagctc	1800
actcattagg	caccccaggc	tttacctttt	atgcttcggg	ctcgtatggt	gtgtggaatt	1860
gtgagcggat	aacaatttca	cacaggaaac	agctatgacc	atgattacgc	caagcttgca	1920
tgcttcagag	tcgactctag	aggatccccg	ggtaccgagc	tctgagcaac	aagttgtaaa	1980
ctcactggct	gataaccctg	atatccaata	tgaagtgtta	actaaccatg	gtgctaacga	2040
aggtcttgcg	tgccaagata	tgggcgcaga	atgggcttct	tgtaacaaaag	taaacatgac	2100
actgtttaac	caaggtgaag	ctgttgactc	aaaagattgg	gctattttact	tccacagcat	2160
tcgtctgatt	ctggagtgtg	acaacgagca	gttcaaaaatc	tctcgtgtaa	cgggtgacct	2220
acataagcta	gaaccaacag	ataagtttga	cggttcgct	gcgggtgaag	aggttgttct	2280
tccattgggt	ggtgaatact	ggcaactatt	tgaaactgac	ttcatgccgg	gtgcattcgt	2340
ttctgctcca	aacgcagaac	ctaagatgat	tgcttctcta	aatactgaag	atgttgcgtc	2400
ttttgtgacg	ggtcttgaag	gtaacaacct	aaaacgtaca	ccagatgaca	acaatgtatt	2460
tgcaaacgct	gtgtctcggt	ttgagaaaaa	cgaagaccta	gcaacacaag	acgtatcaac	2520
cacgttacta	ccaacaccaa	tgacggttga	agcgggtaaa	ggcaaagtag	atatcgcgga	2580
tggtattgcg	ctgcctaaag	acgcattcga	tgcgactcag	ttcgacgca	ttcaagatcg	2640
tgcaagaagt	gtaggtgtgg	acgttcgtgg	tgatcttcct	gtaagcatca	ctgttgttcc	2700
tgcaacttc	accggtgaat	tagcaaaatc	tggtgcttac	gaaatgagca	tcaaaggcga	2760
cggatttgtg	attaaagcgt	tcgaccaagc	aggcgctttc	tacgcagtag	aatctatctt	2820
tggccttgga	gatagccaaa	atgctgattc	tctaccacaa	ctgtctatta	aagatgcgcc	2880
tcgttttgat	taccgtgggt	tgatgttgga	tgtggctcgt	aacttcact	ctaaggacgc	2940
aatccttgca	acgctagacc	aaatggcagc	gtacaagatg	aacaaacttc	accttcacct	3000
aaccgatgat	gaaggctggc	gtttagaaat	cccgggtctg	cctgagctga	cagaagtggg	3060
tgctaaccgt	gttttcgata	cacaagagaa	aagctgttta	ctgcctcagc	ttggctctgg	3120
tccaacgaca	gacaactttg	gctctggcta	cttcagcaaa	gcagactacg	tggaaatctt	3180
gaaatacgcg	aaagcacgta	acattgaagt	gattccagaa	atcgatatgc	cagctcacgc	3240

tcgtgcagca	gtagtatcaa	tggaagctcg	ttacgaccgc	ctaattggaag	aaggtaaaga	3300
agctgaagcg	aacgaatacc	gtctgatgga	tcctcaagat	acatcaaacg	taacgacggg	3360
tcagttctac	aataagcaaa	gcttcatcaa	cccatgtatg	gaatcttcaa	ctcgctttgt	3420
tgataagggtg	atttcagaag	tggcagcaat	gcaccaagaa	gctggcgctc	cactaacaac	3480
ttggcacttc	ggtggtgacg	aagcgaagaa	catcaagcta	ggtgctgggt	tccaagacgt	3540
taacgcagaa	gataaaagta	gctggaaagg	cacgattgac	ctgtctaaac	aagacaagcc	3600
gtttgacacg	tctccacaat	gtcagacgct	aatcacagat	ggcacagtca	gtgactttgc	3660
tcacctacca	agccacttcg	cggagaagat	gtcgaagatt	gttgctgaga	aaggcattcc	3720
aaacttccaa	gcttggcaag	atggtttgaa	atacagtgcg	ggcgaaaaag	cggttcgctac	3780
agaaaatact	cgcgtaaaact	tctgggacgt	tctgtactgg	ggcggtactt	cctcagtgta	3840
cgagtgggtct	aagaaaagggt	acgacgtgat	tgtttctaac	ccagattacg	tgtacatgga	3900
tatgccatac	gaagttgacc	cgaagagacg	tggttactac	tggaacacac	gtgcaacgga	3960
tactcgtgaag	atgtttggct	ttgcaccaga	gaacatgcct	caaaacgcag	aaacttctgt	4020
agatcgcgat	ggcaatggct	ttactggtaa	aggtgaaatc	gaagcgaaac	ctttctacgg	4080
tctatctgca	caactttggg	ctgagacagt	acgtaacgac	gagcaatacg	agtacatggg	4140
attccctcgc	gtcctcgcgt	ctgctcagcg	tgcattggac	cgtgctgact	gggaaaacga	4200
ctacaaagtt	ggtgttgagt	actcgcaaaa	ctctaatact	gttgataaaag	catcgctaaa	4260
ccaagactac	aaccgctttg	cgaacgtact	tgggtcaacgt	gaactggcta	agctagaaaa	4320
atcaggtatt	gactaccgcc	taccagtacc	aggtgcaaaa	gtagaagatg	gtaagctagc	4380
aatgaacggt	cagttccctg	gcgtaacgct	tcaatactct	ctggatgggt	agaactgggt	4440
gacttatgca	gacaacgctc	gtccaaatgt	aactggtgaa	gtcttcatcc	gctcgggtatc	4500
tgcgacaggt	gagaaggtaa	gccgtatcac	tagcgtgaaa	taatagcgct	cagtattcac	4560
taaaatcata	gttccttact	caaagccctc	aacttatggt	gggggctttg	tttatttttc	4620
ttcgaaaaat	aagcgtgatc	ccccgggggc	gcctaccttt	cacgagttgc	gcagtttgct	4680
tgcaagactc	tatgagaagc	agataagcga	taagtttgct	caacatcttc	tcgggcataa	4740
gtcggacacc	atggcatcac	agtatcgtga	tgacagaggc	agggagtggg	acaaaattga	4800
aatcaaataa	tgattttatt	ttgactgata	gtgacctgtt	cggttgcaaca	aattgataag	4860
caatgctttt	ttataatgcc	aacttagtat	aaaaaagctg	aacgagaaac	gtaaaatgat	4920
ataaatatca	atatattaaa	ttagattttg	cataaaaaac	agactacata	atactgtaaa	4980
acacaacata	tgagtcactc	atgaatcaac	tacttagatg	gtattagtga	cctgtaacag	5040
agcattagcg	caaggtgatt	tttgtcttct	tgcgctaatt	ttttgtcatc	aaacctgtcg	5100
catgatcatg	gggctgcagg	aattcgatgg	tcgaatttgc	tttcgaattt	ctgccattca	5160
tccgcttatt	atcacttatt	caggcgtagc	accaggcggt	taagggcacc	aataactgcc	5220
ttaaaaaaat	tacgccccgc	cctgccactc	atcgcagtac	tgttgtaatt	cattaagcat	5280
tctgccgaca	tggaagccat	cacagacggc	atgatgaacc	tgaatcgcca	gcggcatcag	5340
caccttgctg	ccttgcgat	aatatttgcc	catggtgaaa	acgggggcga	agaagttgtc	5400
catattggcc	acgtttaaat	caaaactggg	gaaactcacc	cagggtattg	ctgagacgaa	5460
aaacatatcc	tcaataaacc	ctttagggaa	ataggccagg	ttttaccggt	aacacgccac	5520
atcttgcgaa	tatatgtgta	gaaactgccg	gaaatcgctg	tggtattcac	tccagagcga	5580
tgaaaacggt	tcagtttgct	catggaaaac	gggtgaacaa	gggtgaacac	tatcccatat	5640
caccagctca	ccgtctttca	ttgccatacg				5670

<210> 15  
 <211> 238  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> LacZ(/chitobiase Fusion  
  
 <220>  
 <221> CDS  
 <222> (119)...(240)  
  
 <221> -35\_signal  
 <222> (44)...(50)  
 <223> Lac promoter



[illegible]

<212> DNA  
 <213> Artificial Sequence

<220>  
 <221> CDS  
 <222> (1)...(2643)

<223> dnaA/Chitobiase Fusion

<400> 18

gtg	tca	ctt	tcg	ctt	tgg	cag	cag	tgt	ctt	gcc	cga	ttg	cag	gat	gag	48
Met	Ser	Leu	Ser	Leu	Trp	Gln	Gln	Cys	Leu	Ala	Arg	Leu	Gln	Asp	Glu	
1				5					10					15		

ttg	gta	ccg	agc	tct	gag	caa	caa	gtt	gta	aac	tca	ctg	gct	gat	aac	96
Leu	Val	Pro	Ser	Ser	Glu	Gln	Gln	Val	Val	Asn	Ser	Leu	Ala	Asp	Asn	
			20					25					30			

ctt	gat	atc	caa	tat	gaa	gtg	tta	act	aac	cat	ggg	gct	aac	gaa	ggg	144
Leu	Asp	Ile	Gln	Tyr	Glu	Val	Leu	Thr	Asn	His	Gly	Ala	Asn	Glu	Gly	
		35					40					45				

ctt	gcg	tgc	caa	gat	atg	ggc	gca	gaa	tgg	gct	tct	tgt	aac	aaa	gta	192
Leu	Ala	Cys	Gln	Asp	Met	Gly	Ala	Glu	Trp	Ala	Ser	Cys	Asn	Lys	Val	
	50					55					60					

aac	atg	acg	ctt	gtt	aac	caa	ggg	gaa	gct	gtt	gac	tca	aaa	gat	tgg	240
Asn	Met	Thr	Leu	Val	Asn	Gln	Gly	Glu	Ala	Val	Asp	Ser	Lys	Asp	Trp	
65					70				75					80		

gct	att	tac	ttc	cac	agc	att	cgt	ctg	att	ctg	gat	gtt	gac	aac	gag	288
Ala	Ile	Tyr	Phe	His	Ser	Ile	Arg	Leu	Ile	Leu	Asp	Val	Asp	Asn	Glu	
				85					90					95		

cag	ttc	aaa	atc	tct	cgt	gta	acg	ggg	gac	cta	cat	aag	cta	gaa	cca	336
Gln	Phe	Lys	Ile	Ser	Arg	Val	Thr	Gly	Asp	Leu	His	Lys	Leu	Glu	Pro	
			100					105					110			

aca	gat	aag	ttt	gac	ggc	ttc	gct	gcc	ggg	gaa	gag	gtt	gtt	ctt	cca	384
Thr	Asp	Lys	Phe	Asp	Gly	Phe	Ala	Ala	Gly	Glu	Glu	Val	Val	Leu	Pro	
		115					120						125			

ttg	gtt	ggg	gaa	tac	tgg	caa	cta	ttt	gaa	act	gac	ttc	atg	ccg	ggg	432
Leu	Val	Gly	Glu	Tyr	Trp	Gln	Leu	Phe	Glu	Thr	Asp	Phe	Met	Pro	Gly	
	130					135					140					

gca	ttc	gtt	tct	gct	cca	aac	gca	gaa	cct	aag	atg	att	gct	tct	cta	480
Ala	Phe	Val	Ser	Ala	Pro	Asn	Ala	Glu	Pro	Lys	Met	Ile	Ala	Ser	Leu	
145					150					155				160		

aat	act	gaa	gat	gtt	gcg	tct	ttt	gtg	acg	ggg	ctt	gaa	ggg	aac	aac	528
Asn	Thr	Glu	Asp	Val	Ala	Ser	Phe	Val	Thr	Gly	Leu	Glu	Gly	Asn	Asn	
				165					170					175		

cta	aaa	cgt	aca	cca	gat	gac	aac	aat	gta	ttt	gca	aac	gct	gtg	tct	576
Leu	Lys	Arg	Thr	Pro	Asp	Asp	Asn	Asn	Val	Phe	Ala	Asn	Ala	Val	Ser	
			180					185					190			

cgt ttt gag aaa aac gaa gac cta gca aca caa gac gta tca acc acg	624
Arg Phe Glu Lys Asn Glu Asp Leu Ala Thr Gln Asp Val Ser Thr Thr	
195 200 205	
tta cta cca aca cca atg cac gtt gaa gcg ggt aaa ggc aaa gta gat	672
Leu Leu Pro Thr Pro Met His Val Glu Ala Gly Lys Gly Lys Val Asp	
210 215 220	
atc gcg gat ggt att gcg ctg cct aaa gac gca ttc gat gcg act cag	720
Ile Ala Asp Gly Ile Ala Leu Pro Lys Asp Ala Phe Asp Ala Thr Gln	
225 230 235 240	
ttc gca gcg att caa gat cgt gca gaa gtg gta ggt gtg gac gtt cgt	768
Phe Ala Ala Ile Gln Asp Arg Ala Glu Val Val Gly Val Asp Val Arg	
245 250 255	
ggt gat ctt cct gta agc atc act gtt gtt cct gca gac ttc acc ggt	816
Gly Asp Leu Pro Val Ser Ile Thr Val Val Pro Ala Asp Phe Thr Gly	
260 265 270	
gaa tta gca aaa tct ggt gct tac gaa atg agc atc aaa ggc gac ggt	864
Glu Leu Ala Lys Ser Gly Ala Tyr Glu Met Ser Ile Lys Gly Asp Gly	
275 280 285	
att gtg att aaa gcg ttc gac caa gca ggc gct ttc tac gca gta caa	912
Ile Val Ile Lys Ala Phe Asp Gln Ala Gly Ala Phe Tyr Ala Val Gln	
290 295 300	
tct atc ttt ggc ctg gta gat agc caa aat gct gat tct cta cca caa	960
Ser Ile Phe Gly Leu Val Asp Ser Gln Asn Ala Asp Ser Leu Pro Gln	
305 310 315 320	
ctg tct att aaa gat gcg cct cgt ttt gat tac cgt ggt gtg atg gtg	1008
Leu Ser Ile Lys Asp Ala Pro Arg Phe Asp Tyr Arg Gly Val Met Val	
325 330 335	
gat gtg gct cgt aac ttc cac tct aag gac gca atc ctt gca acg cta	1056
Asp Val Ala Arg Asn Phe His Ser Lys Asp Ala Ile Leu Ala Thr Leu	
340 345 350	
gac caa atg gca gcg tac aag atg aac aaa ctt cac ctt cac cta acc	1104
Asp Gln Met Ala Ala Tyr Lys Met Asn Lys Leu His Leu His Leu Thr	
355 360 365	
gat gat gaa ggc tgg cgt tta gaa atc ccg ggt ctg cct gag ctg aca	1152
Asp Asp Glu Gly Trp Arg Leu Glu Ile Pro Gly Leu Pro Glu Leu Thr	
370 375 380	
gaa gtg ggt gct aac cgt tgt ttc gat aca caa gag aaa agc tgt tta	1200
Glu Val Gly Ala Asn Arg Cys Phe Asp Thr Gln Glu Lys Ser Cys Leu	
385 390 395 400	
ctg cct cag ctt ggc tct ggt cca acg aca gac aac ttt ggc tct ggc	1248
Leu Pro Gln Leu Gly Ser Gly Pro Thr Thr Asp Asn Phe Gly Ser Gly	
405 410 415	

tac ttc agc aaa gca gac tac gtg gaa atc ttg aaa tac gcg aaa gca Tyr Phe Ser Lys Ala Asp Tyr Val Glu Ile Leu Lys Tyr Ala Lys Ala 420 425 430	1296
cgt aac att gaa gtg att cca gaa atc gat atg cca gct cac gct cgt Arg Asn Ile Glu Val Ile Pro Glu Ile Asp Met Pro Ala His Ala Arg 435 440 445	1344
gca gca gta gta tca atg gaa gct cgt tac gac cgc cta atg gaa gaa Ala Ala Val Val Ser Met Glu Ala Arg Tyr Asp Arg Leu Met Glu Glu 450 455 460	1392
ggc aaa gaa gct gaa gcg aac gaa tac cgt ctg atg gat cct caa gat Gly Lys Glu Ala Glu Ala Asn Glu Tyr Arg Leu Met Asp Pro Gln Asp 465 470 475 480	1440
aca tca aac gta acg acg gtt cag ttc tac aat aag caa agc ttc atc Thr Ser Asn Val Thr Thr Val Gln Phe Tyr Asn Lys Gln Ser Phe Ile 485 490 495	1488
aac cca tgt atg gaa tct tca act cgc ttt gtt gat aag gtg att tca Asn Pro Cys Met Glu Ser Ser Thr Arg Phe Val Asp Lys Val Ile Ser 500 505 510	1536
gaa gtg gca gca atg cac caa gaa gct ggc gct cca cta aca act tgg Glu Val Ala Ala Met His Gln Glu Ala Gly Ala Pro Leu Thr Thr Trp 515 520 525	1584
cac ttc ggt ggt gac gaa gcg aag aac atc aag cta ggt gct ggt ttc His Phe Gly Gly Asp Glu Ala Lys Asn Ile Lys Leu Gly Ala Gly Phe 530 535 540	1632
caa gac gtt aac gca gaa gat aaa gta agc tgg aaa ggc acg att gac Gln Asp Val Asn Ala Glu Asp Lys Val Ser Trp Lys Gly Thr Ile Asp 545 550 555 560	1680
ctg tct aaa caa gac aag ccg ttt gca cag tct cca caa tgt cag acg Leu Ser Lys Gln Asp Lys Pro Phe Ala Gln Ser Pro Gln Cys Gln Thr 565 570 575	1728
cta atc aca gat ggc aca gtc agt gac ttt gct cac cta cca agc cac Leu Ile Thr Asp Gly Thr Val Ser Asp Phe Ala His Leu Pro Ser His 580 585 590	1776
ttc gcg gaa gaa gtg tcg aag att gtt gct gag aaa ggc att cca aac Phe Ala Glu Glu Val Ser Lys Ile Val Ala Glu Lys Gly Ile Pro Asn 595 600 605	1824
ttc caa gct tgg caa gat ggt ttg aaa tac agt gac ggc gaa aaa gcg Phe Gln Ala Trp Gln Asp Gly Leu Lys Tyr Ser Asp Gly Glu Lys Ala 610 615 620	1872
ttc gct aca gaa aat act cgc gta aac ttc tgg gac gtt ctg tac tgg Phe Ala Thr Glu Asn Thr Arg Val Asn Phe Trp Asp Val Leu Tyr Trp 625 630 635 640	1920
ggc ggt act tcc tca gtg tac gag tgg tct aag aaa ggt tac gac gtg	1968

Gly Gly Thr Ser Ser Val Tyr Glu Trp Ser Lys Lys Gly Tyr Asp Val	
645 650 655	
att gtt tct aac cca gat tac gtg tac atg gat atg cca tac gaa gtt	2016
Ile Val Ser Asn Pro Asp Tyr Val Tyr Met Asp Met Pro Tyr Glu Val	
660 665 670	
gac ccg aaa gag cgt ggt tac tac tgg gca aca cgt gca acg gat act	2064
Asp Pro Lys Glu Arg Gly Tyr Tyr Trp Ala Thr Arg Ala Thr Asp Thr	
675 680 685	
cgt aag atg ttt ggc ttt gca cca gag aac atg cct caa aac gca gaa	2112
Arg Lys Met Phe Gly Phe Ala Pro Glu Asn Met Pro Gln Asn Ala Glu	
690 695 700	
act tct gta gat cgc gat ggc aat ggc ttt act ggt aaa ggt gaa atc	2160
Thr Ser Val Asp Arg Asp Gly Asn Gly Phe Thr Gly Lys Gly Glu Ile	
705 710 715 720	
gaa gcg aaa cct ttc tac ggt cta tct gca caa ctt tgg tct gag aca	2208
Glu Ala Lys Pro Phe Tyr Gly Leu Ser Ala Gln Leu Trp Ser Glu Thr	
725 730 735	
gta cgt aac gac gag caa tac gag tac atg gta ttc cct cgc gtc ctc	2256
Val Arg Asn Asp Glu Gln Tyr Glu Tyr Met Val Phe Pro Arg Val Leu	
740 745 750	
gct gct gct cag cgt gca tgg cac cgt gct gac tgg gaa aac gac tac	2304
Ala Ala Ala Gln Arg Ala Trp His Arg Ala Asp Trp Glu Asn Asp Tyr	
755 760 765	
aaa gtt ggt gtt gag tac tcg caa aac tct aat cta gtt gat aaa gca	2352
Lys Val Gly Val Glu Tyr Ser Gln Asn Ser Asn Leu Val Asp Lys Ala	
770 775 780	
tcg cta aac caa gac tac aac cgc ttt gcg aac gta ctt ggt caa cgt	2400
Ser Leu Asn Gln Asp Tyr Asn Arg Phe Ala Asn Val Leu Gly Gln Arg	
785 790 795 800	
gaa ctg gct aag cta gaa aaa tca ggt att gac tac cgc cta cca gta	2448
Glu Leu Ala Lys Leu Glu Lys Ser Gly Ile Asp Tyr Arg Leu Pro Val	
805 810 815	
cca ggt gca aaa gta gaa gat ggt aag cta gca atg aac gtt cag ttc	2496
Pro Gly Ala Lys Val Glu Asp Gly Lys Leu Ala Met Asn Val Gln Phe	
820 825 830	
cct ggc gta acg ctt caa tac tct ctg gat ggt gag aac tgg ttg act	2544
Pro Gly Val Thr Leu Gln Tyr Ser Leu Asp Gly Glu Asn Trp Leu Thr	
835 840 845	
tat gca gac aac gct cgt cca aat gta act ggt gaa gtc ttc atc cgc	2592
Tyr Ala Asp Asn Ala Arg Pro Asn Val Thr Gly Glu Val Phe Ile Arg	
850 855 860	
tcg gta tct gcg aca ggt gag aag gta agc cgt atc act agc gtg aaa	2640
Ser Val Ser Ala Thr Gly Glu Lys Val Ser Arg Ile Thr Ser Val Lys	

865 870 875 880

taa  
\*

2643

<210> 19  
<211> 880  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> dnaA/Chitobiase Fusion

<400> 19

Met	Ser	Leu	Ser	Leu	Trp	Gln	Gln	Cys	Leu	Ala	Arg	Leu	Gln	Asp	Glu
1				5					10					15	
Leu	Val	Pro	Ser	Ser	Glu	Gln	Gln	Val	Val	Asn	Ser	Leu	Ala	Asp	Asn
			20					25					30		
Leu	Asp	Ile	Gln	Tyr	Glu	Val	Leu	Thr	Asn	His	Gly	Ala	Asn	Glu	Gly
		35					40					45			
Leu	Ala	Cys	Gln	Asp	Met	Gly	Ala	Glu	Trp	Ala	Ser	Cys	Asn	Lys	Val
	50					55					60				
Asn	Met	Thr	Leu	Val	Asn	Gln	Gly	Glu	Ala	Val	Asp	Ser	Lys	Asp	Trp
65					70					75					80
Ala	Ile	Tyr	Phe	His	Ser	Ile	Arg	Leu	Ile	Leu	Asp	Val	Asp	Asn	Glu
				85					90					95	
Gln	Phe	Lys	Ile	Ser	Arg	Val	Thr	Gly	Asp	Leu	His	Lys	Leu	Glu	Pro
			100					105					110		
Thr	Asp	Lys	Phe	Asp	Gly	Phe	Ala	Ala	Gly	Glu	Glu	Val	Val	Leu	Pro
		115					120					125			
Leu	Val	Gly	Glu	Tyr	Trp	Gln	Leu	Phe	Glu	Thr	Asp	Phe	Met	Pro	Gly
	130					135					140				
Ala	Phe	Val	Ser	Ala	Pro	Asn	Ala	Glu	Pro	Lys	Met	Ile	Ala	Ser	Leu
145					150					155					160
Asn	Thr	Glu	Asp	Val	Ala	Ser	Phe	Val	Thr	Gly	Leu	Glu	Gly	Asn	Asn
			165						170					175	
Leu	Lys	Arg	Thr	Pro	Asp	Asp	Asn	Asn	Val	Phe	Ala	Asn	Ala	Val	Ser
			180				185						190		
Arg	Phe	Glu	Lys	Asn	Glu	Asp	Leu	Ala	Thr	Gln	Asp	Val	Ser	Thr	Thr
		195					200					205			
Leu	Leu	Pro	Thr	Pro	Met	His	Val	Glu	Ala	Gly	Lys	Gly	Lys	Val	Asp
	210					215					220				
Ile	Ala	Asp	Gly	Ile	Ala	Leu	Pro	Lys	Asp	Ala	Phe	Asp	Ala	Thr	Gln
225					230					235					240
Phe	Ala	Ala	Ile	Gln	Asp	Arg	Ala	Glu	Val	Val	Gly	Val	Asp	Val	Arg
				245					250					255	
Gly	Asp	Leu	Pro	Val	Ser	Ile	Thr	Val	Val	Pro	Ala	Asp	Phe	Thr	Gly
		260						265					270		
Glu	Leu	Ala	Lys	Ser	Gly	Ala	Tyr	Glu	Met	Ser	Ile	Lys	Gly	Asp	Gly
	275						280					285			
Ile	Val	Ile	Lys	Ala	Phe	Asp	Gln	Ala	Gly	Ala	Phe	Tyr	Ala	Val	Gln
	290					295					300				
Ser	Ile	Phe	Gly	Leu	Val	Asp	Ser	Gln	Asn	Ala	Asp	Ser	Leu	Pro	Gln
305					310					315					320
Leu	Ser	Ile	Lys	Asp	Ala	Pro	Arg	Phe	Asp	Tyr	Arg	Gly	Val	Met	Val
				325					330					335	
Asp	Val	Ala	Arg	Asn	Phe	His	Ser	Lys	Asp	Ala	Ile	Leu	Ala	Thr	Leu

Asp	Gln	Met	340	Ala	Ala	Tyr	Lys	Met	345	Asn	Lys	Leu	His	Leu	His	Leu	Thr
		355							360					365			
Asp	Asp	Glu	Gly	Trp	Arg	Leu	Glu	Ile	Pro	Gly	Leu	Pro	Glu	Leu	Thr		
	370					375					380						
Glu	Val	Gly	Ala	Asn	Arg	Cys	Phe	Asp	Thr	Gln	Glu	Lys	Ser	Cys	Leu		
385					390					395					400		
Leu	Pro	Gln	Leu	Gly	Ser	Gly	Pro	Thr	Thr	Asp	Asn	Phe	Gly	Ser	Gly		
			405						410					415			
Tyr	Phe	Ser	Lys	Ala	Asp	Tyr	Val	Glu	Ile	Leu	Lys	Tyr	Ala	Lys	Ala		
			420					425					430				
Arg	Asn	Ile	Glu	Val	Ile	Pro	Glu	Ile	Asp	Met	Pro	Ala	His	Ala	Arg		
	435						440					445					
Ala	Ala	Val	Val	Ser	Met	Glu	Ala	Arg	Tyr	Asp	Arg	Leu	Met	Glu	Glu		
	450					455				460							
Gly	Lys	Glu	Ala	Glu	Ala	Asn	Glu	Tyr	Arg	Leu	Met	Asp	Pro	Gln	Asp		
465					470					475					480		
Thr	Ser	Asn	Val	Thr	Thr	Val	Gln	Phe	Tyr	Asn	Lys	Gln	Ser	Phe	Ile		
			485						490					495			
Asn	Pro	Cys	Met	Glu	Ser	Ser	Thr	Arg	Phe	Val	Asp	Lys	Val	Ile	Ser		
		500						505					510				
Glu	Val	Ala	Ala	Met	His	Gln	Glu	Ala	Gly	Ala	Pro	Leu	Thr	Thr	Trp		
	515						520					525					
His	Phe	Gly	Gly	Asp	Glu	Ala	Lys	Asn	Ile	Lys	Leu	Gly	Ala	Gly	Phe		
	530				535						540						
Gln	Asp	Val	Asn	Ala	Glu	Asp	Lys	Val	Ser	Trp	Lys	Gly	Thr	Ile	Asp		
545					550					555					560		
Leu	Ser	Lys	Gln	Asp	Lys	Pro	Phe	Ala	Gln	Ser	Pro	Gln	Cys	Gln	Thr		
			565						570					575			
Leu	Ile	Thr	Asp	Gly	Thr	Val	Ser	Asp	Phe	Ala	His	Leu	Pro	Ser	His		
			580					585					590				
Phe	Ala	Glu	Glu	Val	Ser	Lys	Ile	Val	Ala	Glu	Lys	Gly	Ile	Pro	Asn		
	595					600						605					
Phe	Gln	Ala	Trp	Gln	Asp	Gly	Leu	Lys	Tyr	Ser	Asp	Gly	Glu	Lys	Ala		
	610					615					620						
Phe	Ala	Thr	Glu	Asn	Thr	Arg	Val	Asn	Phe	Trp	Asp	Val	Leu	Tyr	Trp		
625					630					635					640		
Gly	Gly	Thr	Ser	Ser	Val	Tyr	Glu	Trp	Ser	Lys	Lys	Gly	Tyr	Asp	Val		
			645						650					655			
Ile	Val	Ser	Asn	Pro	Asp	Tyr	Val	Tyr	Met	Asp	Met	Pro	Tyr	Glu	Val		
			660					665					670				
Asp	Pro	Lys	Glu	Arg	Gly	Tyr	Tyr	Trp	Ala	Thr	Arg	Ala	Thr	Asp	Thr		
	675						680					685					
Arg	Lys	Met	Phe	Gly	Phe	Ala	Pro	Glu	Asn	Met	Pro	Gln	Asn	Ala	Glu		
	690					695					700						
Thr	Ser	Val	Asp	Arg	Asp	Gly	Asn	Gly	Phe	Thr	Gly	Lys	Gly	Glu	Ile		
705					710					715					720		
Glu	Ala	Lys	Pro	Phe	Tyr	Gly	Leu	Ser	Ala	Gln	Leu	Trp	Ser	Glu	Thr		
			725						730					735			
Val	Arg	Asn	Asp	Glu	Gln	Tyr	Glu	Tyr	Met	Val	Phe	Pro	Arg	Val	Leu		
			740					745					750				
Ala	Ala	Ala	Gln	Arg	Ala	Trp	His	Arg	Ala	Asp	Trp	Glu	Asn	Asp	Tyr		
	755						760					765					
Lys	Val	Gly	Val	Glu	Tyr	Ser	Gln	Asn	Ser	Asn	Leu	Val	Asp	Lys	Ala		
	770					775					780						
Ser	Leu	Asn	Gln	Asp	Tyr	Asn	Arg	Phe	Ala	Asn	Val	Leu	Gly	Gln	Arg		
785					790					795					800		

Glu	Leu	Ala	Lys	Leu	Glu	Lys	Ser	Gly	Ile	Asp	Tyr	Arg	Leu	Pro	Val	
				805					810						815	
Pro	Gly	Ala	Lys	Val	Glu	Asp	Gly	Lys	Leu	Ala	Met	Asn	Val	Gln	Phe	
			820					825						830		
Pro	Gly	Val	Thr	Leu	Gln	Tyr	Ser	Leu	Asp	Gly	Glu	Asn	Trp	Leu	Thr	
			835					840					845			
Tyr	Ala	Asp	Asn	Ala	Arg	Pro	Asn	Val	Thr	Gly	Glu	Val	Phe	Ile	Arg	
	850					855					860					
Ser	Val	Ser	Ala	Thr	Gly	Glu	Lys	Val	Ser	Arg	Ile	Thr	Ser	Val	Lys	
865					870					875					880	

18

1

002030"160E960